

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : <b>C12N 15/49, C07K 14/16</b>		A2	(11) International Publication Number: <b>WO 00/39303</b> (43) International Publication Date: <b>6 July 2000 (06.07.00)</b>
(21) International Application Number: <b>PCT/US99/31272</b> (22) International Filing Date: <b>30 December 1999 (30.12.99)</b>		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(30) Priority Data: 60/114,495 31 December 1998 (31.12.98) US 60/156,670 29 September 1999 (29.09.99) US		(71) Applicant: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).	
(72) Inventors: BARNETT, Susan; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US). HARTOG, Karin; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US). MARTIN, Eric; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US).		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(74) Agents: DOLLARD, Anne, S.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US) et al.			
(54) Title: <b>MODIFIED HIV ENV POLYPEPTIDES</b>			
(57) Abstract			
Polynucleotide encoding modified HIV Env polypeptides are disclosed. The Env polypeptides are modified so as to expose at least part of the CD4 binding region. Methods of diagnosis, treatment and prevention using the polynucleotides and polypeptides are also provided.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

MODIFIED HIV ENV POLYPEPTIDESTechnical Field

5 The invention relates generally to modified HIV envelope (Env) polypeptides which are useful as immunizing agents or for generating an immune response in a subject, for example a cellular immune response or a protective immune response. More particularly, the invention relates Env polypeptides such as gp120, gp140 or gp160, wherein at least one of the native  $\beta$ -sheet configurations has been modified. The invention also pertains to methods 10 of using these polypeptides to elicit an immune response against a broad range of HIV subtypes.

Background of the Invention

15 The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders. (see, e.g., Barre-Sinoussi, et al., (1983) *Science* 220:868-871; Gallo et al. (1984) *Science* 224:500-503; Levy et al., (1984) *Science* 225:840-842; Siegal et al., (1981) *N. Engl. J. Med.* 305:1439-1444). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous 20 system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury et al., (1986) *Science* 231:850-853). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner et al., (1985) *Nature* 313:277-284; Sanchez-Pescador et al., (1985) *Science* 227:484-492), including three structural genes -- gag, pol and env -- that are common to all 25 retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses, SIV (previously referred to as STLV-III), also contain these structural genes. (Guyader et al., (1987) *Nature* 326:662-669; Chakrabarti et al., (1987) *Nature*

30 The envelope protein of HIV-1, HIV-2 and SIV is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in the

membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. gp120 and gp41 are more covalently associated and free gp120 can be released from the surface of virions and infected cells.

As depicted in Figure 1, crystallography studies of the gp120 core polypeptide 5 indicate that this polypeptide is folded into two major domains having certain emanating structures. The inner domain (inner with respect to the N and C terminus) features a two-helix, two-stranded bundle with a small five-stranded  $\beta$ -sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem emanates. The outer domain is a staked double barrel that lies along side the inner domain so that the outer barrel and inner 10 bundle axes are approximately parallel. Between the distal inner domain and the distal outer domain is a four-stranded bridging sheet which holds a peculiar minidomain in contact with, but distinct from, the inner, the outer domain, and the V1/V2 domain. The bridging sheet is composed of four  $\beta$ -strand structures ( $\beta$ -3,  $\beta$ -2,  $\beta$ -21,  $\beta$ -20, shown in Figure 1). The bridging region can be seen in Figure 1 packing primarily over the inner domain, although some 15 surface residues of the outer domain, such as Phe 382, reach into the bridging sheet to form part of its hydrophobic core.

The basic unit of the  $\beta$ -sheet conformation of the bridging sheet region is the  $\beta$ -strand which exists as a less tightly coiled helix, with 2.0 residues per turn. The  $\beta$ -strand conformation is only stable when incorporated into a  $\beta$ -sheet, where hydrogen bonds with 20 close to optimal geometry are formed between the peptide groups on adjacent  $\beta$ -strands; the dipole moments of the strands are also aligned favorably. Side chains from adjacent residues of the same strand protrude from opposite sides of the sheet and do not interact with each other, but have significant interactions with their backbone and with the side chains of neighboring strands. For a general description of  $\beta$ -sheets, see, e.g., T.E. Creighton, Proteins: 25 Structures and Molecular Properties (W.H. Freeman and Company, 1993); and A.L. Lehninger, Biochemistry (Worth Publishers, Inc., 1975).

The gp120 polypeptide is instrumental in mediating entry into the host cell. Recent studies have indicated that binding of CD4 to gp120 induces a conformational change in Env that allows for binding to a co-receptor (e.g. a chemokine receptor) and subsequent entry of 30 the virus into the cell. (Wyatt, R., et al. (1998) *Nature* 393:705-711; Kwong, P., et al. (1998) *Nature* 393:648-659). Referring again to Figure 1, CD4 is bound into a depression formed at the interface of the outer domain, the inner domain and the bridging sheet of gp120.

Immunogenicity of the gp120 polypeptide has also been studied. For example, individuals infected by HIV-1 usually develop antibodies that can neutralize the virus in *in vitro* assays, and this response is directed primarily against linear neutralizing determinants in the third variable loop of gp120 glycoprotein (Javaherian, K., et al. (1989) *Proc. Natl. Acad. Sci.* 86:6786-6772; Matsushita, M., et al. (1988) *J. Virol.* 62:2107-2144; Putney, S., et al. (1986) *Science* 234:1392-1395; Rushe, J. R., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85: 3198-3202.). However, these antibodies generally exhibit the ability to neutralize only a limited number of HIV-1 strains (Matthews, T. (1986) *Proc. Natl. Acad. Sci. USA* 83:9709-9713; Nara, P. L., et al. (1988) *J. Virol.* 62:2622-2628; Parker, T. J., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:1932-1936). Later in the course of HIV infection in humans, antibodies capable of neutralizing a wider range of HIV-1 isolates appear (Barre-Sinoussi, F., et al. (1983) *Science* 220:868-871; Robert-Guroff, M., et al. (1985) *Nature* (London) 316:72-74; Weis, R., et al. (1985) *Nature* (London) 316:69-72; Weis, R., et al. (1986) *Nature* (London) 324:572-575).

Recent work done by Stamatatos et al (1998) *AIDS Res Hum Retroviruses* 14(13):1129-39, shows that a deletion of the variable region 2 from a HIV-1<sub>SF162</sub> virus, which utilizes the CCR-5 co-receptor for virus entry, rendered the virus highly susceptible to serum-mediated neutralization. This V2 deleted virus was also neutralized by sera obtained from patients infected not only with clade B HIV-1 isolates but also with clade A, C, D and F HIV-1 isolates. However, deletion of the variable region 1 had no effect. Deletion of the variable regions 1 and 2 from a LAI isolate HIV-1<sub>IIIb</sub> also increased the susceptibility to neutralization by monoclonal antibodies whose epitopes are located within the V3 loop, the CD4-binding site, and conserved gp120 regions (Wyatt, R., et al. (1995) *J Virol.* 69:5723-5733). Rabbit immunogenicity studies done with the HIV-1 virus with deletions in the V1/V2 and V3 region from the LAI strain, which uses the CXCR4 co-receptor for virus entry, showed no improvement in the ability of Env to raise neutralizing antibodies (Leu et al. (1998) *AIDS Res. and Human Retroviruses*. 14:151-155).

Further, a subset of the broadly reactive antibodies, found in most infected individuals, interferes with the binding of gp120 and CD4 (Kang, C.-Y., et al. (1991) *Proc. Natl. Acad. Sci. USA*. 88:6171-6175; McDougal, J. S., et al. (1986) *J. Immunol.* 137:2937-2944). Other antibodies are believed to bind to the chemokine receptor binding region after CD4 has bound to Env (Thali et al. (1993) *J. Virol.* 67:3978-3988). The fact that neutralizing

antibodies generated during the course of HIV infection do not provide permanent antiviral effect may in part be due to the generation of "neutralization escapes" virus mutants and to the general decline in the host immune system associated with pathogenesis. In contrast, the presence of pre-existing neutralizing antibodies upon initial HIV-1 exposure will likely have 5 a protective effect.

It is widely thought that a successful vaccine should be able to induce a strong, broadly neutralizing antibody response against diverse HIV-1 strains (Montefiori and Evans 10 (1999) *AIDS Res. Hum. Ret.* 15(8):689-698; Bolognesi, D.,P., et al. (1994) *Ann. Int. Med.* 8:603-611; Haynes, B., F., et al. (1996) *Science* ;271: 324-328.). Neutralizing antibodies, by attaching to the incoming virions, can reduce or even prevent their infectivity for target cells and prevent the cell-to-cell spread of virus in tissue culture (Hu et al. (1992) *Science* 255:456-459; Burton, D.,R. and Montefiori, D. (1997) *AIDS* 11(suppl. A): 587-598). However as described above, antibodies directed against gp120 do not generally exhibit broad antibody responses against different HIV strains.

15 Currently, the focus of vaccine development, from the perspective of humoral immunity, is on the neutralization of primary isolates that utilize the CCR5 chemokine co-receptor believed to be important in virus entry (Zhu, T., et al. (1993) *Science* 261:1179-1181; Fiore, J., et al. (1994) *Virology*; 204:297-303). These viruses are generally much more 20 resistant to antibody neutralization than T-cell line adapted strains that use the CXCR4 co-receptor, although both can be neutralized *in vitro* by certain broadly and potent acting monoclonal antibodies, such as IgG1b12, 2G12 and 2F5 (Trkola, A., et al. (1995) *J. Virol.* 69:6609-6617; D'Sousa PM., et al (1997) *J. Infect. Dis.* 175:1062-1075). These monoclonal 25 antibodies are directed to the CD4 binding site, a glycosylation site and to the gp41 fusion domain, respectively. The problem that remains, however, is that it is not known how to induce antibodies of the appropriate specificity by vaccination. Antibodies (Abs) elicited by gp120 glycoprotein from a given isolate are usually only able to neutralize closely related viruses generally from similar, usually from the same, HIV-1 subtype.

30 Despite the above approaches, there remains a need for Env antigens that can elicit an immunological response (e.g., neutralizing and/or protective antibodies) in a subject against multiple HIV strains and subtypes, for example when administered as a vaccine. The present invention solves these and other problems by providing modified Env polypeptides (e.g., gp120) to expose epitopes in or near the CD4 binding site.

Summary of the Invention

In accordance with the present invention, modified HIV Env polypeptides are provided. In particular, deletions and/or mutations are made in one or more of the 4- $\beta$  antiparallel-bridging sheet in the HIV Env polypeptide. In this way, enough structure is left 5 to allow correct folding of the polypeptide, for example of gp120, yet enough of the bridging sheet is removed to expose the CD4 groove, allowing an immune response to be generated against epitopes in or near the CD4 binding site of the Env polypeptide (e.g., gp120).

In one aspect, the invention includes a polynucleotide encoding a modified HIV Env 10 polypeptide wherein the polypeptide has at least one modified (e.g., deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example the constructs depicted in Figures 6-29 (SEQ ID NOs:3 to 26). In certain embodiments, the polynucleotide also has the region corresponding to residues 124-198 of the polypeptide HXB-2 (e.g., V1/V2) deleted and at least one amino acid deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210, relative to 15 HXB-2. In other embodiments, these polynucleotides encode Env polypeptides having at least one amino acid of the small loop of the bridging sheet (e.g., amino acid residues 427 to 429 relative to HXB-2) deleted or replaced. The amino acid sequences of the modified polypeptides encoded by the polynucleotides of the present invention can be based on any HIV variant, for example SF162.

20 In another aspect, the invention includes immunogenic modified HIV Env polypeptides having at least one modified (e.g., deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example a deletion or replacement of one amino acids in the small loop region (e.g., amino acid residues 25 427 to 429 relative to HXB-2). These polypeptides may have modifications (e.g., a deletion or a replacement) of at least one amino acid between about amino acid residue 420 and amino acid residue 436, relative to HXB-2 and, optionally, may have deletions or truncations of the V1 and/or V2 regions. The immunogenic, modified polypeptides of the present invention can be based on any HIV variant, for example SF162.

30 In another aspect, the invention includes a vaccine composition comprising any of the polynucleotides encoding modified Env polypeptides described above. Vaccine compositions comprising the modified Env polypeptides and, optionally, an adjuvant are also included in the invention.

In yet another aspect, the invention includes a method of inducing an immune response in subject comprising, administering one or more of the polynucleotides or constructs described above in an amount sufficient to induce an immune response in the subject. In certain embodiments, the method further comprises administering an adjuvant to the subject.

5 In another aspect, the invention includes a method of inducing an immune response in a subject comprising administering a composition comprising any of the modified Env polypeptides described above and an adjuvant. The composition is administered in an amount sufficient to induce an immune response in the subject.

10 In another aspect, the invention includes a method of inducing an immune response in a subject comprising

(a) administering a first composition comprising any of the polynucleotides described above in a priming step and

15 (b) administering a second composition comprising any of the modified Env polypeptides described above, as a booster, in an amount sufficient to induce an immune response in the subject. In certain embodiments, the first composition, the second composition or both the first and second compositions further comprise an adjuvant.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

20

#### Brief Description of the Drawings

Figure 1 is a schematic depiction of the tertiary structure of the HIV-1<sub>HXB-2</sub> Env gp120 polypeptide, as determined by crystallography studies.

25 Figures 2A-C depict alignment of the amino acid sequence of wild-type HIV-1<sub>HXB-2</sub> Env gp160 polypeptide (SEQ ID NO:1) with amino acid sequence of HIV variants SF162 (shown as "162") (SEQ ID NO:2), SF2, CM236 and US4. Arrows indicate the regions that are deleted or replaced in the modified polypeptides. Black dots indicate conserved cysteine residues. The star indicates the position of the last amino acid in gp120.

30 Figures 3A-J depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having V1/V2 deletions. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 4A-M depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

5 Figures 5A-N depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having both V1/V2 deletions and, in addition, deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figure 6 depicts the nucleotide sequence of the construct designated Val120-Ala204  
10 (SEQ ID NO:3).

Figure 7 depicts the nucleotide sequence of the construct designated Val120-Ile201  
(SEQ ID NO:4).

Figure 8 depicts the nucleotide sequence of the construct designated Val120-Ile201B  
(SEQ ID NO:5).

15 Figure 9 depicts the nucleotide sequence of the construct designated Lys121-Val200  
(SEQ ID NO:6).

Figure 10 depicts the nucleotide sequence of the construct designated Leu122-Ser199  
(SEQ ID NO:7).

20 Figure 11 depicts the nucleotide sequence of the construct designated Val120-Thr202  
(SEQ ID NO:8).

Figure 12 depicts the nucleotide sequence of the construct designated Trp427-Gly431  
(SEQ ID NO:9).

Figure 13 depicts the nucleotide sequence of the construct designated Arg426-Gly431  
(SEQ ID NO:10).

25 Figure 14 depicts the nucleotide sequence of the construct designated Arg426-  
Gly431B (SEQ ID NO:11).

Figure 15 depicts the nucleotide sequence of the construct designated Arg426-Lys432  
(SEQ ID NO:12).

30 Figure 16 depicts the nucleotide sequence of the construct designated Asn425-Lys432  
(SEQ ID NO:13).

Figure 17 depicts the nucleotide sequence of the construct designated Ile424-Ala433  
(SEQ ID NO:14).

Figure 18 depicts the nucleotide sequence of the construct designated Ile423-Met434 (SEQ ID NO:15).

Figure 19 depicts the nucleotide sequence of the construct designated Gln422-Tyr435 (SEQ ID NO:16).

5 Figure 20 depicts the nucleotide sequence of the construct designated Gln422-Tyr435B (SEQ ID NO:17).

Figure 21 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Gly431 (SEQ ID NO:18).

10 Figure 22 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Lys432 (SEQ ID NO:19).

Figure 23 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Trp427-Gly431 (SEQ ID NO:20).

Figure 24 depicts the nucleotide sequence of the construct designated Lys121-Val200;Asn425-Lys432 (SEQ ID NO:21).

15 Figure 25 depicts the nucleotide sequence of the construct designated Val120-Ile201;Ile424-Ala433 (SEQ ID NO:22).

Figure 26 depicts the nucleotide sequence of the construct designated Val120-Ile201B; Ile424-Ala433 (SEQ ID NO:23).

20 Figure 27 depicts the nucleotide sequence of the construct designated Val120-Thr202;Ile424-Ala433 (SEQ ID NO:24).

Figure 28 depicts the nucleotide sequence of the construct designated Val127-Asn195 (SEQ ID NO:25).

Figure 29 depicts the nucleotide sequence of the construct designated Val127-Asn195; Arg426-Gly431 (SEQ ID NO:26).

25

#### Detailed Description of the Invention

The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, viral immunobiology, molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., T.E. Creighton, Proteins: Structures and Molecular Properties (W.H. Freeman and Company, 1993); Nelson L.M. and Jerome H.K. HIV Protocols in Methods in Molecular Medicine, vol. 17, 1999; Sambrook, et al., Molecular Cloning: A

Laboratory Manual (Cold Spring Harbor Laboratory, 1989); F.M. Ausubel et al. Current Protocols in Molecular Biology, Greene Publishing Associates & Wiley Interscience New York; and Lipkowitz and Boyd, Reviews in Computational Chemistry, volumes 1-present (Wiley-VCH, New York, New York, 1999).

5 It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a polypeptide" includes a mixture of two or more polypeptides, and the like.

## 10 Definitions

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

15 The terms "polypeptide," and "protein" are used interchangeably herein to denote any polymer of amino acid residues. The terms encompass peptides, oligopeptides, dimers, multimers, and the like. Such polypeptides can be derived from natural sources or can be synthesized or recombinantly produced. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc.

20 A polypeptide as defined herein is generally made up of the 20 natural amino acids Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y) and Val (V) and may also include any of the several known amino acid analogs, both naturally occurring and synthesized analogs, such as but not limited to homoisoleucine, asaleucine, 2-(methylenecyclopropyl)glycine, S-methylcysteine, S-(prop-1-enyl)cysteine, homoserine, ornithine, norleucine, norvaline, homoarginine, 3-(3-carboxyphenyl)alanine, 25 cyclohexylalanine, mimosine, pipecolic acid, 4-methylglutamic acid, canavanine, 2,3-diaminopropionic acid, and the like. Further examples of polypeptide agents which will find use in the present invention are set forth below.

30 By "geometry" or "tertiary structure" of a polypeptide or protein is meant the overall 3-D configuration of the protein. As described herein, the geometry can be determined, for example, by crystallography studies or by using various programs or algorithms which predict the geometry based on interactions between the amino acids making up the primary and secondary structures.

By "wild type" polypeptide, polypeptide agent or polypeptide drug, is meant a naturally occurring polypeptide sequence, and its corresponding secondary structure. An "isolated" or "purified" protein or polypeptide is a protein which is separate and discrete from a whole organism with which the protein is normally associated in nature. It is apparent that 5 the term denotes proteins of various levels of purity. Typically, a composition containing a purified protein will be one in which at least about 35%, preferably at least about 40-50%, more preferably, at least about 75-85%, and most preferably at least about 90% or more, of the total protein in the composition will be the protein in question.

By "Env polypeptide" is meant a molecule derived from an envelope protein, 10 preferably from HIV Env. The envelope protein of HIV-1 is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in (and spans) the membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. As there is no covalent attachment between gp120 and gp41, free 15 gp120 is released from the surface of virions and infected cells. Env polypeptides may also include gp140 polypeptides. Env polypeptides can exist as monomers, dimers or multimers.

By a "gp120 polypeptide" is meant a molecule derived from a gp120 region of the Env polypeptide. Preferably, the gp120 polypeptide is derived from HIV Env. The primary 20 amino acid sequence of gp120 is approximately 511 amino acids, with a polypeptide core of about 60,000 daltons. The polypeptide is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five 25 hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence of the HIV-1<sub>HXB-2</sub> (hereinafter "HXB-2") strain, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to most, if not all, gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Despite this variation, most, if not all, gp120 sequences preserve the virus's ability to bind to the viral receptor CD4. A "gp120 polypeptide" includes both single subunits or multimers.

30 Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel  $\beta$ -strands ( $\beta$ -2,  $\beta$ -3,  $\beta$ -20 and  $\beta$ -21) that form a  $\beta$ -sheet. Extruding from one pair of the  $\beta$ -strands ( $\beta$ -2 and  $\beta$ -3) are two loops, V1 and V2. The  $\beta$ -2

sheet occurs at approximately amino acid residue 119 (Cys) to amino acid residue 123 (Thr) while  $\beta$ -3 occurs at approximately amino acid residue 199 (Ser) to amino acid residue 201 (Ile), relative to HXB-2. The "V1/V2 region" occurs at approximately amino acid positions 126 (Cys) to residue 196 (Cys), relative to HXB-2. (see, e.g., Wyatt et al. (1995) *J. Virol.* 69:5723-5733; Stamatatos et al. (1998) *J. Virol.* 72:7840-7845). Extruding from the second pair of  $\beta$ -strands ( $\beta$ -20 and  $\beta$ -21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." In HXB-2,  $\beta$ -20 extends from about amino acid residue 422 (Gln) to amino acid residue 426 (Met) while  $\beta$ -21 extends from about amino acid residue 430 (Val) to amino acid residue 435 (Tyr). In variant SF162, the Met-426 is an Arg (R) residue. The "small loop" extends from about amino acid residue 427 (Trp) through 429 (Lys), relative to HXB-2. A representative diagram of gp120 showing the bridging sheet, the small loop, and V1/V2 is shown in Figure 1. In addition, alignment of the amino acid sequences of Env polypeptide gp160 of selected variants is shown, relative to HXB-2, in Figures 2A-C.

Furthermore, an "Env polypeptide" or "gp120 polypeptide" as defined herein is not limited to a polypeptide having the exact sequence described herein. Indeed, the HIV genome is in a state of constant flux and contains several variable domains which exhibit relatively high degrees of variability between isolates. It is readily apparent that the terms encompass Env (e.g., gp120) polypeptides from any of the identified HIV isolates, as well as newly identified isolates, and subtypes of these isolates. Descriptions of structural features are given herein with reference to HXB-2. One of ordinary skill in the art in view of the teachings of the present disclosure and the art can determine corresponding regions in other HIV variants (e.g., isolates HIV<sub>IIIb</sub>, HIV<sub>SF2</sub>, HIV-1<sub>SF162</sub>, HIV-1<sub>SF170</sub>, HIV<sub>LAV</sub>, HIV<sub>LAI</sub>, HIV<sub>MN</sub>, HIV-1<sub>CM235</sub>, HIV-1<sub>US4</sub>, other HIV-1 strains from diverse subtypes (e.g., subtypes, A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2<sub>UC1</sub> and HIV-2<sub>UC2</sub>), and simian immunodeficiency virus (SIV). (See, e.g., *Virology*, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); *Virology*, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify  $\beta$ -sheet regions). The actual amino acid sequences of the modified Env polypeptides can be based on any HIV variant.

Additionally, the term "Env polypeptide" (e.g., "gp120 polypeptide") encompasses proteins which include additional modifications to the native sequence, such as additional internal deletions, additions and substitutions. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through naturally occurring 5 mutational events. Thus, for example, if the Env polypeptide is to be used in vaccine compositions, the modifications must be such that immunological activity (i.e., the ability to elicit an antibody response to the polypeptide) is not lost. Similarly, if the polypeptides are to be used for diagnostic purposes, such capability must be retained.

Thus, a "modified Env polypeptide" is an Env polypeptide (e.g., gp120 as defined 10 above), which has been manipulated to delete or replace all or a part of the bridging sheet portion and, optionally, the variable regions V1 and V2. Generally, modified Env (e.g., gp120) polypeptides have enough of the bridging sheet removed to expose the CD4 binding site, but leave enough of the structure to allow correct folding (e.g., correct geometry). Thus, modifications to the  $\beta$ -20 and  $\beta$ -21 regions (between about amino acid residues 420 and 435 15 relative to HXB-2) are preferred. Additionally, modifications to the  $\beta$ -2 and  $\beta$ -3 regions (between about amino acid residues 119 (Cys) and 201 (Ile)) and modifications (e.g., truncations) to the V1 and V2 loop regions may also be made. Although not all possible  $\beta$ -sheet and V1/V2 modifications have been exemplified herein, it is to be understood that other disrupting modifications are also encompassed by the present invention.

20 Normally, such a modified polypeptide is capable of secretion into growth medium in which an organism expressing the protein is cultured. However, for purposes of the present invention, such polypeptides may also be recovered intracellularly. Secretion into growth media is readily determined using a number of detection techniques, including, e.g., polyacrylamide gel electrophoresis and the like, and immunological techniques such as 25 Western blotting and immunoprecipitation assays as described in, e.g., International Publication No. WO 96/04301, published February 15, 1996.

A gp120 or other Env polypeptide is produced "intracellularly" when it is found 30 within the cell, either associated with components of the cell, such as in association with the endoplasmic reticulum (ER) or the Golgi Apparatus, or when it is present in the soluble cellular fraction. The gp120 and other Env polypeptides of the present invention may also be secreted into growth medium so long as sufficient amounts of the polypeptides remain

present within the cell such that they can be purified from cell lysates using techniques described herein.

An "immunogenic" gp120 or other Env protein is a molecule that includes at least one epitope such that the molecule is capable of either eliciting an immunological reaction in an 5 individual to which the protein is administered or, in the diagnostic context, is capable of reacting with antibodies directed against the HIV in question.

By "epitope" is meant a site on an antigen to which specific B cells and/or T cells respond, rendering the molecule including such an epitope capable of eliciting an 10 immunological reaction or capable of reacting with HIV antibodies present in a biological sample. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site." An epitope can comprise 3 or more amino acids in a spatial conformation unique to the epitope. Generally, an epitope consists of at least 5 such amino acids and, more 15 usually, consists of at least 8-10 such amino acids. Methods of determining spatial conformation of amino acids are known in the art and include, for example, x-ray crystallography and 2-dimensional nuclear magnetic resonance. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art, such as by the use of hydrophobicity studies and by site-directed serology. See, also, Geysen et al., *Proc. Natl. Acad. Sci. USA* (1984) 81:3998-4002 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given 20 antigen); U.S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., *Molecular Immunology* (1986) 23:709-715 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay showing the ability of one antibody to block the binding of another antibody to a target antigen.

25 An "immunological response" or "immune response" as used herein is the development in the subject of a humoral and/or a cellular immune response to the Env (e.g., gp120) polypeptide when the polypeptide is present in a vaccine composition. These antibodies may also neutralize infectivity, and/or mediate antibody-complement or antibody dependent cell cytotoxicity to provide protection to an immunized host. Immunological 30 reactivity may be determined in standard immunoassays, such as a competition assays, well known in the art.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH

package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: <http://www.ncbi.nlm.gov/cgi-bin/BLAST>.

One of skill in the art can readily determine the proper search parameters to use for a given sequence in the above programs. For example, the search parameters may vary based on the size of the sequence in question. Thus, for example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about 50% identity to Y contiguous nucleotides derived from any of the sequences described herein, (ii) X equals Y, and (iii) X is greater than or equal to 6 nucleotides and up to 5000 nucleotides, preferably greater than or equal to 8 nucleotides and up to 5000 nucleotides, more preferably 10-12 nucleotides and up to 5000 nucleotides, and even more preferably 15-20 nucleotides, up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing and claims), including all integer values falling within the above-described ranges.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% sequence (including all integer values falling within these described ranges) identity to the synthetic expression cassette sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as the query sequence.

Computer programs are also available to determine the likelihood of certain polypeptides to form structures such as  $\beta$ -sheets. One such program, described herein, is the "ALB" program for protein and polypeptide secondary structure calculation and predication. In addition, secondary protein structure can be predicted from the primary amino acid sequence, for example using protein crystal structure and aligning the protein sequence related to the crystal structure (e.g., using Molecular Operating Environment (MOE) programs available from the Chemical Computing Group Inc., Montreal, P.Q., Canada). Other methods of predicting secondary structures are described, for example, in Garnier et al. (1996) *Methods Enzymol.* 266:540-553; Geourjon et al. (1995) *Comput. Applic. Biosci.* 11:681-684; Levin (1997) *Protein Eng.* 10:771-776; and Rost et al. (1993) *J. Molec. Biol.* 232:584-599.

Homology can also be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. 15 Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. DNA 20 sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., *supra*; *DNA Cloning, supra*; *Nucleic Acid Hybridization, supra*.

A "coding sequence" or a sequence which "encodes" a selected protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to cDNA from viral nucleotide sequences as well as 25 synthetic and semisynthetic DNA sequences and sequences including base analogs. A transcription termination sequence may be located 3' to the coding sequence.

"Control elements" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control elements need always be present so long as the desired gene is capable of being transcribed and translated.

5 A control element "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Operably linked" refers to an arrangement of elements wherein the components so 10 described are configured so as to perform their usual function. Thus, control elements operably linked to a coding sequence are capable of effecting the expression of the coding sequence when RNA polymerase is present. The control elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between, e.g., a 15 promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a 20 polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such 25 terms denoting prokaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the 30 parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

By "vertebrate subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including 5 rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated 10 from an individual, including but not limited to, for example, blood, plasma, serum, fecal matter, urine, bone marrow, bile, spinal fluid, lymph fluid, samples of the skin, external secretions of the skin, respiratory, intestinal, and genitourinary tracts, samples derived from the gastric epithelium and gastric mucosa, tears, saliva, milk, blood cells, organs, biopsies and also samples of *in vitro* cell culture constituents including but not limited to conditioned 15 media resulting from the growth of cells and tissues in culture medium, e.g., recombinant cells, and cell components.

The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemiluminescers, enzymes, 20 enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof which is capable of exhibiting fluorescence in the detectable range. Particular examples of labels which may be used with the invention include, but are not limited to fluorescein, rhodamine, dansyl, umbelliferone, Texas red, luminol, acridinium 25 esters, NADPH,  $\alpha$ - $\beta$ -galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase and urease.

### Overview

The present invention concerns modified Env polypeptide molecules (e.g., 30 glycoprotein ("gp") 120). Without being bound by a particular theory, it appears that it has been difficult to generate immunological responses against Env because the CD4 binding site is buried between the outer domain, the inner domain and the V1/V2 domains. Thus, although deletion of the V1/V2 domain may render the virus more susceptible to

neutralization by monoclonal antibody directed to the CD4 site, the bridging sheet covering most of the CD4 binding domain may prevent an antibody response. Thus, the present invention provides Env polypeptides that maintain their general overall structure yet expose the CD4 binding domain. This allows the generation of an immune response (e.g., an antibody response) to epitopes in or near the CD4 binding site.

5 Various forms of the different embodiments of the invention, described herein, may be combined.

### **β-Sheet Conformations**

10 In the present invention, location of the β-sheet structures were identified relative to 3-D (crystal) structure of an HXB-2 crystallized Env protein (see, Example 1A). Based on this structure, constructs encoding polypeptides having replacements and or excisions which maintain overall geometry while exposing the CD4 binding site were designed. In particular, the crystal structure of HXB-2 was downloaded from the Brookhaven Database. Using the 15 default parameters of the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package, homology and fit of amino acids which could replace the native loops between β-strands yet maintain overall tertiary structure were determined. Constructs encoding the modified Env polypeptides were then designed (Example 1.B.).

Thus, the modified Env polypeptides typically have enough of the bridging sheet 20 removed to expose the CD4 groove, but have enough of the structure to allow correct folding of the Env glycoprotein. Exemplary constructs are described below.

### **Polypeptide Production**

25 The polypeptides of the present invention can be produced in any number of ways which are well known in the art.

In one embodiment, the polypeptides are generated using recombinant techniques, well known in the art. In this regard, oligonucleotide probes can be devised based on the known sequences of the Env (e.g., gp120) polypeptide genome and used to probe genomic or cDNA libraries for Env genes. The gene can then be further isolated using standard 30 techniques and, e.g., restriction enzymes employed to truncate the gene at desired portions of the full-length sequence. Similarly, the Env gene(s) can be isolated directly from cells and tissues containing the same, using known techniques, such as phenol extraction and the

sequence further manipulated to produce the desired truncations. *See, e.g.,* Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA.

The genes encoding the modified (e.g., truncated and/or substituted) polypeptides can be produced synthetically, based on the known sequences. The nucleotide sequence can be 5 designed with the appropriate codons for the particular amino acid sequence desired. The complete sequence is generally assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. *See, e.g.,* Edge (1981) *Nature* 292:756; Nambair et al. (1984) *Science* 223:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311; Stemmer et al. (1995) *Gene* 164:49-53.

10 Recombinant techniques are readily used to clone a gene encoding an Env polypeptide gene which can then be mutagenized *in vitro* by the replacement of the appropriate base pair(s) to result in the codon for the desired amino acid. Such a change can include as little as one base pair, effecting a change in a single amino acid, or, can encompass several base pair changes. Alternatively, the mutations can be effected using a mismatched 15 primer which hybridizes to the parent nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by keeping the mutant base centrally located. *See, e.g.,* Innis et al, (1990) PCR Applications: Protocols for Functional Genomics; Zoller and Smith, 20 *Methods Enzymol.* (1983) 100:468. Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization probe. The technique is also applicable for generating multiple point mutations. *See, e.g.,* Dalbie-McFarland et al. *Proc. Natl. Acad. Sci USA* (1982) 79:6409.

25 Once coding sequences for the desired proteins have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. As will be apparent from the teachings herein, a wide variety of vectors encoding modified polypeptides can be generated by creating expression constructs which operably link, in various combinations, polynucleotides encoding Env polypeptides having deletions or mutation therein. Thus, 30 polynucleotides encoding a particular deleted V1/V2 region can be operably linked with polynucleotides encoding polypeptides having deletions or replacements in the small loop

region and the construct introduced into a host cell for polypeptide expression. Non-limiting examples of such combinations are discussed in the Examples.

Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage  $\lambda$  (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), YCp19 (*Saccharomyces*) and 10 bovine papilloma virus (mammalian cells). *See, generally, DNA Cloning: Vols. I & II, supra; Sambrook et al., supra; B. Perbal, supra.*

Insect cell expression systems, such as baculovirus systems, can also be used and are known to those of skill in the art and described in, e.g., Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987). Materials and methods for 15 baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit).

Plant expression systems can also be used to produce the modified Env proteins. Generally, such systems use virus-based vectors to transfect plant cells with heterologous genes. For a description of such systems see, e.g., Porta et al., *Mol. Biotech.* (1996) 5:209-221; and Hackland et al., *Arch. Virol.* (1994) 139:1-22.

Viral systems, such as a vaccinia based infection/transfection system, as described in Tomei et al., *J. Virol.* (1993) 67:4017-4026 and Selby et al., *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first transfected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA 25 polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high 30 level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired Env polypeptide is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The 5 coding sequence may or may not contain a signal peptide or leader sequence. With the present invention, both the naturally occurring signal peptides or heterologous sequences can be used. Leader sequences can be removed by the host in post-translational processing. *See, e.g.*, U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Such sequences include, but are not limited to, the TPA leader, as well as the honey bee mellitin signal sequence.

10 Other regulatory sequences may also be desirable which allow for regulation of expression of the protein sequences relative to the growth of the host cell. Such regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may 15 also be present in the vector, for example, enhancer sequences.

The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

20 In some cases it may be necessary to modify the coding sequence so that it may be attached to the control sequences with the appropriate orientation; *i.e.*, to maintain the proper reading frame. Mutants or analogs may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or 25 more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. *See, e.g.*, Sambrook *et al.*, *supra*; *DNA Cloning*, Vols. I and II, *supra*; *Nucleic Acid Hybridization*, *supra*.

The expression vector is then used to transform an appropriate host cell. A number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), such as, but not limited to, Chinese hamster 30 ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (*e.g.*, Hep G2), Vero293 cells, as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find

use with the present expression constructs. Yeast hosts useful in the present invention include *inter alia*, *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guillermondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Insect cells for use with baculovirus expression vectors include, *inter alia*, *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni*.

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. The selection of the appropriate growth conditions is within the skill of the art.

In one embodiment, the transformed cells secrete the polypeptide product into the surrounding media. Certain regulatory sequences can be included in the vector to enhance secretion of the protein product, for example using a tissue plasminogen activator (TPA) leader sequence, a  $\gamma$ -interferon signal sequence or other signal peptide sequences from known secretory proteins. The secreted polypeptide product can then be isolated by various techniques described herein, for example, using standard purification techniques such as but not limited to, hydroxyapatite resins, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoabsorbent techniques, affinity chromatography, immunoprecipitation, and the like..

Alternatively, the transformed cells are disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the Env polypeptides substantially intact. Intracellular proteins can also be obtained by removing components from the cell wall or membrane, e.g., by the use of detergents or organic solvents, such that leakage of the Env polypeptides occurs. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990)

For example, methods of disrupting cells for use with the present invention include but are not limited to: sonication or ultrasonication; agitation; liquid or solid extrusion; heat treatment; freeze-thaw; desiccation; explosive decompression; osmotic shock; treatment with lytic enzymes including proteases such as trypsin, neuraminidase and lysozyme; alkali treatment; and the use of detergents and solvents such as bile salts, sodium dodecylsulphate,

Triton, NP40 and CHAPS. The particular technique used to disrupt the cells is largely a matter of choice and will depend on the cell type in which the polypeptide is expressed, culture conditions and any pre-treatment used.

Following disruption of the cells, cellular debris is removed, generally by 5 centrifugation, and the intracellularly produced Env polypeptides are further purified, using standard purification techniques such as but not limited to, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoabsorbent techniques, affinity chromatography, immunoprecipitation, and the like.

For example, one method for obtaining the intracellular Env polypeptides of the 10 present invention involves affinity purification, such as by immunoaffinity chromatography using anti-Env specific antibodies, or by lectin affinity chromatography. Particularly preferred lectin resins are those that recognize mannose moieties such as but not limited to resins derived from *Galanthus nivalis* agglutinin (GNA), *Lens culinaris* agglutinin (LCA or lentil lectin), *Pisum sativum* agglutinin (PSA or pea lectin), *Narcissus pseudonarcissus* agglutinin (NPA) and *Allium ursinum* agglutinin (AUA). The choice of a suitable affinity 15 resin is within the skill in the art. After affinity purification, the Env polypeptides can be further purified using conventional techniques well known in the art, such as by any of the techniques described above.

It may be desirable to produce Env (e.g., gp120) complexes, either with itself or other 20 proteins. Such complexes are readily produced by e.g., co-transfected host cells with constructs encoding for the Env (e.g., gp120) and/or other polypeptides of the desired complex. Co-transfection can be accomplished either in *trans* or *cis*, i.e., by using separate vectors or by using a single vector which bears both of the Env and other gene. If done using a single vector, both genes can be driven by a single set of control elements or, alternatively, 25 the genes can be present on the vector in individual expression cassettes, driven by individual control elements. Following expression, the proteins will spontaneously associate. Alternatively, the complexes can be formed by mixing the individual proteins together which have been produced separately, either in purified or semi-purified form, or even by mixing culture media in which host cells expressing the proteins, have been cultured. See, 30 International Publication No. WO 96/04301, published February 15, 1996, for a description of such complexes.

Relatively small polypeptides, i.e., up to about 50 amino acids in length, can be conveniently synthesized chemically, for example by any of several techniques that are known to those skilled in the peptide art. In general, these methods employ the sequential addition of one or more amino acids to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid is protected by a suitable protecting group. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions that allow for the formation of an amide linkage. The protecting group is then removed from the newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support, if solid phase synthesis techniques are used) are removed sequentially or concurrently, to render the final polypeptide. By simple modification of this general procedure, it is possible to add more than one amino acid at a time to a growing chain, for example, by coupling (under conditions which do not racemize chiral centers) a protected tripeptide with a properly protected dipeptide to form, after deprotection, a pentapeptide. See, e.g., J. M. Stewart and J. D. Young, Solid Phase Peptide Synthesis (Pierce Chemical Co., Rockford, IL 1984) and G. Barany and R. B. Merrifield, The Peptides: Analysis, Synthesis, Biology, editors E. Gross and J. Meienhofer, Vol. 2, (Academic Press, New York, 1980), pp. 3-254, for solid phase peptide synthesis techniques; and M. Bodansky, Principles of Peptide Synthesis, (Springer-Verlag, Berlin 1984) and E. Gross and J. Meienhofer, Eds., The Peptides: Analysis, Synthesis, Biology, Vol. 1, for classical solution synthesis.

Typical protecting groups include t-butyloxycarbonyl (Boc), 9-fluorenylmethoxycarbonyl (Fmoc) benzylloxycarbonyl (Cbz); p-toluenesulfonyl (Tx); 2,4-dinitrophenyl; benzyl (Bzl); biphenylisopropylloxycarboxy-carbonyl, t-amylloxycarbonyl, isobornyloxycarbonyl, o-bromobenzylloxycarbonyl, cyclohexyl, isopropyl, acetyl, o-nitrophenylsulfonyl and the like.

Typical solid supports are cross-linked polymeric supports. These can include divinylbenzene cross-linked-styrene-based polymers, for example, divinylbenzene-hydroxymethylstyrene copolymers, divinylbenzene-chloromethylstyrene copolymers and divinylbenzene-benzhydrylaminopolystyrene copolymers.

The polypeptide analogs of the present invention can also be chemically prepared by other methods such as by the method of simultaneous multiple peptide synthesis. See, e.g., Houghten *Proc. Natl. Acad. Sci. USA* (1985) 82:5131-5135; U.S. Patent No. 4,631,211.

5

### Diagnostic and Vaccine Applications

The intracellularly produced Env polypeptides of the present invention, complexes thereof, or the polynucleotides coding therefor, can be used for a number of diagnostic and therapeutic purposes. For example, the proteins and polynucleotides or antibodies generated against the same, can be used in a variety of assays, to determine the presence of reactive 10 antibodies/and or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

The presence of antibodies reactive with the Env (e.g., gp120) polypeptides and, conversely, antigens reactive with antibodies generated thereto, can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as 15 competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays; radioimmunoassays; immunoelectrophoresis; immunoprecipitation, etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, or enzymatic labels or dye molecules, or other 20 methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

Solid supports can be used in the assays such as nitrocellulose, in membrane or microtiter well form; polyvinylchloride, in sheets or microtiter wells; polystyrene latex, in beads or microtiter plates; polyvinylidene fluoride; diazotized paper; nylon membranes; 25 activated beads, and the like.

Typically, the solid support is first reacted with the biological sample (or the gp120 proteins), washed and then the antibodies, (or a sample suspected of containing antibodies), applied. After washing to remove any non-bound ligand, a secondary binder moiety is added under suitable binding conditions, such that the secondary binder is capable of associating 30 selectively with the bound ligand. The presence of the secondary binder can then be detected using techniques well known in the art. Typically, the secondary binder will comprise an antibody directed against the antibody ligands. A number of anti-human immunoglobulin

(Ig) molecules are known in the art (e.g., commercially available goat anti-human Ig or rabbit anti-human Ig). Ig molecules for use herein will preferably be of the IgG or IgA type, however, IgM may also be appropriate in some instances. The Ig molecules can be readily conjugated to a detectable enzyme label, such as horseradish peroxidase, glucose oxidase,

5 Beta-galactosidase, alkaline phosphatase and urease, among others, using methods known to those of skill in the art. An appropriate enzyme substrate is then used to generate a detectable signal.

Alternatively, a "two antibody sandwich" assay can be used to detect the proteins of the present invention. In this technique, the solid support is reacted first with one or more of 10 the antibodies directed against Env (e.g., gp120), washed and then exposed to the test sample. Antibodies are again added and the reaction visualized using either a direct color reaction or using a labeled second antibody, such as an anti-immunoglobulin labeled with horseradish peroxidase, alkaline phosphatase or urease.

Assays can also be conducted in solution, such that the viral proteins and antibodies 15 thereto form complexes under precipitating conditions. The precipitated complexes can then be separated from the test sample, for example, by centrifugation. The reaction mixture can be analyzed to determine the presence or absence of antibody-antigen complexes using any of a number of standard methods, such as those immunodiagnostic methods described above.

The modified Env proteins, produced as described above, or antibodies to the 20 proteins, can be provided in kits, with suitable instructions and other necessary reagents, in order to conduct immunoassays as described above. The kit can also contain, depending on the particular immunoassay used, suitable labels and other packaged reagents and materials (i.e. wash buffers and the like). Standard immunoassays, such as those described above, can be conducted using these kits.

25 The Env polypeptides and polynucleotides encoding the polypeptides can also be used in vaccine compositions, individually or in combination, in e.g., prophylactic (i.e., to prevent infection) or therapeutic (to treat HIV following infection) vaccines. The vaccines can comprise mixtures of one or more of the modified Env proteins (or nucleotide sequences encoding the proteins), such as Env (e.g., gp120) proteins derived from more than one viral 30 isolate. The vaccine may also be administered in conjunction with other antigens and immunoregulatory agents, for example, immunoglobulins, cytokines, lymphokines, and chemokines, including but not limited to IL-2, modified IL-2 (cys125-ser125), GM-CSF, IL-

12,  $\gamma$ -interferon, IP-10, MIP1 $\beta$  and RANTES. The vaccines may be administered as polypeptides or, alternatively, as naked nucleic acid vaccines (e.g., DNA), using viral vectors (e.g., retroviral vectors, adenoviral vectors, adeno-associated viral vectors) or non-viral vectors (e.g., liposomes, particles coated with nucleic acid or protein). The vaccines may also

5 comprise a mixture of protein and nucleic acid, which in turn may be delivered using the same or different vehicles. The vaccine may be given more than once (e.g., a "prime" administration followed by one or more "boosts") to achieve the desired effects. The same composition can be administered as the prime and as the one or more boosts. Alternatively, different compositions can be used for priming and boosting.

10 The vaccines will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

15 A carrier is optionally present which is a molecule that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Furthermore, the Env

20 polypeptide may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc.

25 Adjuvants may also be used to enhance the effectiveness of the vaccines. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y

30 (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size

emulsion, and (c) Ribi<sup>TM</sup> adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox<sup>TM</sup>); (3) saponin adjuvants, such as Stimulon<sup>TM</sup> (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isogluatme (nor-MDP), N-acetylmuramyl-L-alanyl-D-isogluatminyl-L-alanine-2-(l'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Typically, the vaccine compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above.

The vaccines will comprise a therapeutically effective amount of the modified Env proteins, or complexes of the proteins, or nucleotide sequences encoding the same, and any other of the above-mentioned components, as needed. By "therapeutically effective amount" is meant an amount of a modified Env (e.g., gp120) protein which will induce a protective immunological response in the uninfected, infected or unexposed individual to which it is administered. Such a response will generally result in the development in the subject of a secretory, cellular and/or antibody-mediated immune response to the vaccine. Usually, such

a response includes but is not limited to one or more of the following effects; the production of antibodies from any of the immunological classes, such as immunoglobulins A, D, E, G or M; the proliferation of B and T lymphocytes; the provision of activation, growth and differentiation signals to immunological cells; expansion of helper T cell, suppressor T cell, and/or cytotoxic T cell.

5 Preferably, the effective amount is sufficient to bring about treatment or prevention of disease symptoms. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the individual to be treated; the capacity of the individual's immune system to synthesize antibodies; the degree of protection desired; the 10 severity of the condition being treated; the particular Env polypeptide selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. A "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

Once formulated, the nucleic acid vaccines may be accomplished with or without viral 15 vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient 20 presence of DNA in the recipient. Both nucleic acids and/or peptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and 25 vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

While the invention has been described in conjunction with the preferred specific 30 embodiments thereof, it is to be understood that the foregoing description as well as the examples which follow are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

Experimental

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

5 Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

## EXAMPLE 1

10 A.1. Best-Fit and Homology Searches

The crystal structure of HXB-2 gp 120 was downloaded from the Brookhaven database (COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) 15-JUN-98 1GC1 TITLE: HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY). Beta strands 3, 2, 21, and 20 of gp 120 form a sheet near the CD4 binding site. Strands  $\beta$ -3 and  $\beta$ -2 are connected by the V1/V2 loop. Strands  $\beta$ -21 and  $\beta$ -20 are connected by another small loop. The H-bonds at the interface between strands  $\beta$ -2 and  $\beta$ -21 are the only connection between domains of the "lower" half of the protein (joining helix alpha 1 to the CD4 binding site). This beta sheet and these loops mask some antigens (e.g., antigens which may generate neutralizing antibodies) that are only exposed during the 15 CD4 binding.

20 Constructs that remove enough of the beta sheet to expose the antigens in the CD4 binding site, but leave enough of the protein to allow correct folding were designed. Specifically targeted were modifications to the small loop and, optional deletion of the V1/V2 loops. Three different types of constructs were designed: (1) constructs encoding 25 polypeptides that leave the number of residues making up the entire 4-strand beta sheet intact, but replace one or more residues; (2) constructs that encode polypeptide having at least one residue of at least one beta strand excised or (3) constructs encoding polypeptides having at least two residues of at least one beta strand excised. Thus, a total of 6 different turns were needed to rejoin the ends of the strands.

30 Initially, residues in the small loop (residues 427-430, relative to HXB-2) and connected beta strands ( $\beta$ -20 and  $\beta$ -21) were modified to contain Gly and Pro (common in beta turns). These sequences were then used as the target to match in each search. The

geometry of the target was matched to known proteins in the Brookhaven Protein Data Bank. In particular, 5-residue turns (including an overlapping single residue at the N-terminal, the 2 residue target turn and 2 overlapping residues at the C-terminal) were searched in the databases. In other words, these modified loops add a 2 residue turn that should be able to support a geometry that will maintain the beta-sheet structure of the wild type protein. The calculations were performed using the default parameters in the Loop Search feature of the Biopolymer module of the Syby1 molecular modeling package. In each case, the 25 best fits based on geometry alone were reviewed and, of those, several selected for homology and fit.

In addition, it was also determined what modifications could be made to remove most of the V1/V2 loop (residues 124-198, relative to HXB-2) yet leave the geometry of the protein intact. As with the small loop, constructs were also designed which excised one or more residues from the  $\beta$ -2 strand (residues 119-123 of HXB-2), the  $\beta$ -3 strand (residues 199-201 of HXB-2) or both  $\beta$ -2 and  $\beta$ -3. For these constructs, known loops were searched to match the geometry of a pentamer (including two remaining residues from the N-terminal side, a 2 residue turn and 1 C-terminal residue). For these searches, Gly-Gly was preferred as the insert along with at least one C-terminal substitution.

#### A.2. Small Loop Replacements

In one aspect, the native sequence was replaced with residues that expose the CD4 binding site, but leave the overall geometry of the protein relatively unchanged. For the small loop replacements, the target to match was: ASN425-MET426-GLY427-GLY428-GLY431. Results of the search are summarized in Table 1.

Table 1: Search of Small Loop (Asn425 through Gly431)

25

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit	LYS-ASP-SER-ASN-ASN	0.16689	62.5	27
3	TYR-GLY-LEU-GLY-LEU	0.220308	62.5	28
4	GLU-ARG-GLU-ASP-GLY	0.241754	62.5	29
7	ARG-LYS-GLY-GLY-ASN	0.24881	100	30
12	TRP-THR-GLY-SER-TYR	0.26417	83.33	31

Based on these results, constructs encoding Gly-Gly (#7), Gly-Ser (#12) or Gly-Gly-Asn (#7) were recommended.

As V1/V2 and one or more residues of  $\beta$ -2 and  $\beta$ -3 are also optionally deleted in the modified polypeptides of the invention, known loops to match the geometry of the V1/V2 loop were also searched. The V1/V2 loop the target to match was: Lys121-Leu-122-Gly123-Gly124-Ser199. Some notable matches are shown in Table 2:

Table 2: Search of V1/V2 loop (Lys121 through Ser199)

Rank	Sequence	RMSD	% Homology	Seq Id. No.
10	Best fit	GLN-VAL-HIS-ASP-GLU	0.154764	68.75
	2	LYS-GLU-GLY-ASP-LYS	0.15718	81.25
	9	ARG-SER-GLY-ARG-SER	0.173731	68.75
	11	THR-LEU-GLY-ASN-SER	0.175554	81.25
	16	HIS-PHE-GLY-ALA-GLY	0.178772	93.75

15

Based on these searches, constructs encoding Gly-Asn in place of V1/V2 were recommended.

### A.3. One Additional Residue Excisions

20

For a slightly truncated small loop, one more residue was trimmed from each beta strand to slightly shorten the beta sheet. The target to match was: ILE424-ASN425-GLY426-GLY427-LYS432. Results are shown in Table 3:

Table 3: Search of Beta sheet shortened by One residue (Ile424 through Lys432)

Rank	Sequence	RMSD	% Homology	Seq Id No.
25	Best fit:	ARG-MET-ALA-PRO-VAL	0.316805	58.33
	Best hom:	ASP-SER-ASP-GLY-PRO	0.440896	83.33

Although these searches showed more variation and worse fits than the previous truncation, the Pro-Val or Pro-Leu encoding constructs were very similar. Accordingly, Ala-Pro encoding constructs were recommended.

Sequences encoding gp120 polypeptides having V1/V2 deleted and an additional residue from  $\beta$ -2 or  $\beta$ -3 excised were also searched. The V1/V2 loop the target to match was: 5 VAL120-LYS121-GLY122-GLY123-VAL200. Some notable matches are shown in Table 4.

Table 4: Search of V1/V2 loop (Val120 through Val200)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-VAL-ASP-PRO-TYR	0.400892	58.33333	39
2	SER-THR-ASN-PRO-LEU	0.402575	54.16667	40
3	THR-ARG-SER-PRO-LEU	0.403965	58.33333	41
7	ARG-MET-ALA-PRO-VAL	0.440118	58.33333	42

15

The construct encoding Ala-Pro (e.g., #7) was recommended.

#### A.4. Further Excisions

In yet another truncation, an additional residue was trimmed from the  $\beta$ -20 and  $\beta$ -21 strands to further shorten the beta sheet. The target to match was ILE423-ILE424-GLY425-GLY426-ALA433. Notable matches are shown in Table 5.

Table 5: Search of Beta sheet shortened by Two Residues (Ile423 through Ala433)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-TYR-GLU-GLY-VAL	0.130107	79.16666	43
2	GLN-VAL-GLY-ASN-THR	0.138245	79.16666	44
3:	THR-VAL-GLY-GLY-ILE	0.153362	100	45

30 A construct encoding Gly-Gly (e.g., #3), which has 100% homology, was recommended.

Also searched were sequences encoding a deleted V1/V2 region and at least two residues excised from  $\beta$ -2,  $\beta$ -3 or at least one residue excised from  $\beta$ -2 and  $\beta$ -3. The target to match was: CYS119-VAL120-GLY121-GLY122-ILE201. Notable matches are shown in Table 6.

5

Table 6: Search of V1/V2 loop (Cys119 through Ile201)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	ASP-LEU-PRO-GLY-CYS	0.250501	75	46
4	ASP-VAL-GLY-GLY-LEU	0.290383	100	47

10

It was determined that both constructs would be used.

#### B.1. Constructs encoding modified Env polypeptides

As described above, the native loops extruding from the 4- $\beta$  antiparallel-stands were excised and replaced with 1 to 3 residue turns. The loops were replaced so as to leave the entire  $\beta$ -strands or excised by trimming one or more amino acid from each side of the connected strands. The ends of the strands were rejoined with turns that preserve the same backbone geometry (e.g., tertiary structure of  $\beta$ -20 and  $\beta$ -21), as determined by searching the Brookhaven Protein Data Bank.

15

Table 7A is a summary of the truncations of the variable regions 1 and 2

recommended for this study, as determined in Example 1.A. above.

20

Table 7A

V1/V2 Modifications	SEQ ID NO	Figure
-LEU122-GLY-ASN-SER199	7	10
-LYS121-ALA-PRO-VAL200-	6	9
-VAL120-GLY-GLY-ILE201-	4	7
-VAL120-PRO-GLY-ILE201B-	5	8
-VAL120-GLY-ALA-GLY-ALA204-	3	6
-VAL120-GLY-GLY-ALA-THR202-	8	11
-VAL127-GLY-ALA-GLY-ASN195-	25	28

10

As previously noted, the polypeptides encoded by the constructs of the present invention are numbered relative to HXB-2, but the particular amino acid residue of the polypeptides encoded by these exemplary constructs is based on SF-162. Thus, for example, although amino acid residue 195 in HXB-2 is a serine (S), constructs encoding polypeptides having the wild type SF162 sequence will have an asparagine (N) at this position. Table 7B shows just three of the variations in amino acid sequence between strains HXB-2 and SF162. The entire sequences, including differences in residue and amino acid number, of HXB-2 and SF162 are shown in the alignment of Figure 2 (SEQ ID NOs:1 and 2).

20

Table 7B

HXB-2 amino acid number	HXB-2 Residue	SF162 Residue/amino acid number
128	Serine (S)	Thr (T)/114
195	Serine (S)	Asn (N)/188
426	Met (M)	Arg (R)/411

25

Constructs containing deletions in the  $\beta$ -20 strand,  $\beta$ -21 stand and small loop were also constructed. Shown in Table 8 are constructs encoding truncations in these regions. The constructs in Table 8 are numbered relative to HXB-2 but the unmodified amino acid sequence is based on SF162. Thus, the construct encodes an arginine (Arg) as is found in

SF162 in the amino acid position numbered 426 relative to HXB-2 (See, also, Table 7B).

Changes from wildtype (SF162) are shown in bold in Table 8B.

Table 8

Small Loop/β-20 and β-21 (Modified)		SEQ ID NO	Figure
5	-TRP427-GLY-GLY431-	9	12
	-ARG426-GLY-GLY-GLY431-	10	13
	-ARG426-GLY-SER-GLY431B-	11	14
	-ARG426-GLY-GLY-ASN-LYS432-	12	15
10	-ASN425-ALA-PRO-LYS432-	13	16
	-ILE424-GLY-GLY-ALA433-	14	17
	-ILE423-GLY-GLY-MET434-	15	18
	GLN422-GLY-GLY-TYR435-	16	19
15	-GLN422-ALA-PRO-TYR435B-	17	20

The deletion constructs shown in Tables 7 and 8 for each one of the β-strands and combinations of them are constructed. These deletions will be tested in the Env forms gp120, gp140 and gp160 from different HIV strains like subtype B strains (e.g., SF162, US4, SF2), subtype E strains (e.g., CM235) and subtype C strains (e.g., AF110968 or AF110975).

20 Exemplary constructs for SF162 are shown in the Figures and are summarized in Table 9. As noted above in Figure 2 and Table 7B, in the bridging sheet region, the amino acid sequence of SF162 differs from HXB-2 in that the Met426 of HXB-2 is an Arg in SF162. In Table 9, V1/V2 refers to deletions in the V1/V2 region; # bsm refers to a modification in the bridging sheet small loop.

25

Table 9

Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Val120-Ala204	3	6	V1/V2: Val120-Gly-Ala-Gly-Ala204
Val120-Ile201	4	7	V1/V2: Val120-Gly-Gly-Ile201
30	Val120-Ile201B	8	V1/V2: Val120-Pro-Gly-Ile201
	Lys121-Val200	9	V1/V2: Lys121-Ala-Pro-Val200

Table 9

Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Leu122-Ser199	7	10	V1/V2: Leu122-Gly-Asn-Ser199
Val120-Thr202	8	11	V1/V2: Val120-Gly-Gly-Ala-Thr202
Trp427-Gly431	9	12	bsm: Trp427-Gly-Gly431
Arg426-Gly431	10	13	bsm: Arg426-Gly-Gly-Gly431
5 Arg426-Gly431B	11	14	bsm: Arg426-Gly-Ser-Gly431
Arg426-Lys432	12	15	bsm: Arg426-Gly-Gly-Asn-Lys432
Asn425-Lys432	13	16	bsm: Asn425-Ala-Pro-Lys432
Ile424-Ala433	14	17	bsm: Ile424-Gly-Gly-Ala433
Ile423-Met434	15	18	bsm: Ile423-Gly-Gly-Met434
10 Gln422-Tyr435	16	19	bsm: Gln422-Gly-Gly-Tyr435
Val127-Asn195	25	28	bsm: Val127-Gly-Ala-Gly-Asn195
Gln422-Tyr435B	17	20	bsm: Gln422-Ala-Pro-Tyr435
15 Leu122-Ser199; Arg426-Gly431	18	21	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Arg426-Gly-Gly-Gly431
Leu122-Ser199; Arg426-Lys432	19	22	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Arg426-Gly-Gly-Asn-Lys432
Leu122-Ser199-Trp427-Gly431	20	23	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Trp427-Gly-Gly431
20 Lys121-Val200-Asn425-Lys432	21	24	V1/V2/bsm: Lys121-Ala-Pro-Val200 --- Asn425-Ala-Pro-Lys432
Val120-Ile201-Ile424-Ala433	22	25	V1/V2/bsm: Val120-Gly-Gly-Ile201 --- Ile424-Gly-Gly-Ala433
25 Val120-Ile201B-Ile424-Ala433	23	26	V1/V2/bsm: Val120-Pro-Gly-Ile201 --- Ile424-Gly-Gly-Ala433
Val120-Thr202; Ile424-Ala433	24	27	V1/V2/bsm: Val120-Gly-Gly-Ala-Thr202 --- Ile424-Gly-Gly-Ala433
Val127-Asn195; Arg426-Gly431	25	29	V1/V2/bsm: Val127-Gly-Ala-Gly-Asn195 --- Arg426-Gly-Gly-Gly431

30

Combinations of V1/V2 deletions and bridging sheet small loop modifications in addition to those specifically shown in Table 9 are also within the scope of the present invention. Various forms of the different embodiments of the invention, described herein, may be combined.

The first screening will be done after transient expression in COS-7, RD and/or 293 cells. The proteins that are expressed will be analyzed by immunoblot, ELISA, and for binding to mAbs directed to the CD4 binding site and other important epitopes on gp120 to determine integrity of structure. They will also be tested in a CD4 binding assay and, in 5 addition, the binding of neutralizing antibodies, for example using patient sera or mAb 448D (directed to Glu370 and Tyr384, a region of the CD4 binding groove that is not altered by the deletions).

The immunogenicity of these novel Env glycoproteins will be tested in rodents and primates. The structures will be administered as DNA vaccines or adjuvanted protein 10 vaccines or in combined modalities. The goal of these vaccinations will be to archive broadly reactive neutralizing antibody responses.

Claims:

What is claimed is:

- 5        1. A polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one amino acid deleted or replaced in the region corresponding to residues 420 to 436 relative to HXB-2 (SEQ ID NO:1).
- 10        2. The polynucleotide of claim 1, wherein the region corresponding to residues 124-198 relative to HXB-2 is deleted and at least one amino acid is deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210 relative to HXB-2 (SEQ ID NO:1).
- 15        3. The polynucleotide of claim 1, wherein at least one amino acid in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
- 20        4. The polynucleotide of claim 2, wherein at least one amino acid of the in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
- 25        5. The polynucleotide of claim 1, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 30        6. An immunogenic modified HIV Env polypeptide having at least one amino acid deleted or replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
- 35        7. The polypeptide of claim 6, wherein one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

8. The polypeptide of claim 6, wherein more than one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

5 9. The polypeptide of claim 6, wherein at least one amino acid is replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

10 10. The polypeptide of claim 6, wherein at least one amino acid residue between about amino acid residue 427 and amino acid residue 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.

10

11. The polypeptide of claim 6, wherein the V1 and V2 regions of the polypeptide are truncated.

15

12. The polypeptide of claim 10, wherein the V1 and V2 regions of the polypeptide are truncated.

13. The polypeptide of claim 6, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.

20

14. A construct comprising the nucleotide sequence depicted in Figure 6 (SEQ ID NO:3).

15. A construct comprising the nucleotide sequence depicted in Figure 7 (SEQ ID NO:4).

25

16. A construct comprising the nucleotide sequence depicted in Figure 8 (SEQ ID NO:5).

30

17. A construct comprising the nucleotide sequence depicted in Figure 9 (SEQ ID NO:6).

18. A construct comprising the nucleotide sequence depicted in Figure 10 (SEQ ID NO:7).

19. A construct comprising the nucleotide sequence depicted in Figure 11 (SEQ ID 5 NO:8).

20. A construct comprising the nucleotide sequence depicted in Figure 12 (SEQ ID NO:9).

10 21. A construct comprising the nucleotide sequence depicted in Figure 13 (SEQ ID NO:10).

22. A construct comprising the nucleotide sequence depicted in Figure 14 (SEQ ID NO:11).

15 23. A construct comprising the nucleotide sequence depicted in Figure 15 (SEQ ID NO:12).

20 24. A construct comprising the nucleotide sequence depicted in Figure 16 (SEQ ID NO:13).

25 25. A construct comprising the nucleotide sequence depicted in Figure 17 (SEQ ID NO:14).

26. A construct comprising the nucleotide sequence depicted in Figure 18 (SEQ ID NO:15).

27. A construct comprising the nucleotide sequence depicted in Figure 19 (SEQ ID NO:16).

30 28. A construct comprising the nucleotide sequence depicted in Figure 20 (SEQ ID NO:17).

29. A construct comprising the nucleotide sequence depicted in Figure 21 (SEQ ID NO:18).

5 30. A construct comprising the nucleotide sequence depicted in Figure 22 (SEQ ID NO:19).

10 31. A construct comprising the nucleotide sequence depicted in Figure 23 (SEQ ID NO:20).

15 32. A construct comprising the nucleotide sequence depicted in Figure 24 (SEQ ID NO:21).

33. A construct comprising the nucleotide sequence depicted in Figure 25 (SEQ ID NO:22).

15 34. A construct comprising the nucleotide sequence depicted in Figure 26 (SEQ ID NO:23).

20 35. A construct comprising the nucleotide sequence depicted in Figure 27 (SEQ ID NO:24).

36. A construct comprising the nucleotide sequence depicted in Figure 28 (SEQ ID NO:25).

25 37. A construct comprising the nucleotide sequence depicted in Figure 29 (SEQ ID NO:26).

38. A vaccine composition comprising a polynucleotide encoding a modified Env polypeptide according to any one of claims 1-5.

30 39. A vaccine composition comprising a polynucleotide construct encoding a modified Env polypeptide according to any of claims 14-37.

40. A vaccine composition comprising a modified Env polypeptide according to any of claims 6-13.

41. The vaccine composition of any of claims 38-40, further comprising an adjuvant.

5

42. A method of inducing an immune response in subject comprising, administering a polynucleotide according to any one of claims 1-5 in an amount sufficient to induce an immune response in the subject.

10

43. A method of inducing an immune response in subject comprising, administering a polynucleotide construct according to any one of claims 14-37 in an amount sufficient to induce an immune response in the subject.

15

44. A method of inducing an immune response in a subject comprising administering a composition comprising a modified Env polypeptide according to any one of claims 6-13, wherein the composition is administered in an amount sufficient to induce an immune response in the subject

20

45. The method of any of claims 42-44 further comprising administering an adjuvant to the subject.

25

46. A method of inducing an immune response in a subject comprising  
(a) administering a first composition comprising a polynucleotide according to any of claims 1-5 in a priming step and

(b) administering a second composition comprising a modified Env polypeptide according to any of claims 6-13, as a booster, in an amount sufficient to induce an immune response in the subject.

30

47. The method of claim 46 wherein the first composition or second composition further comprise an adjuvant.

48. The method of claim 46 wherein the first and second compositions further comprise an adjuvant.

# gp120 core structure

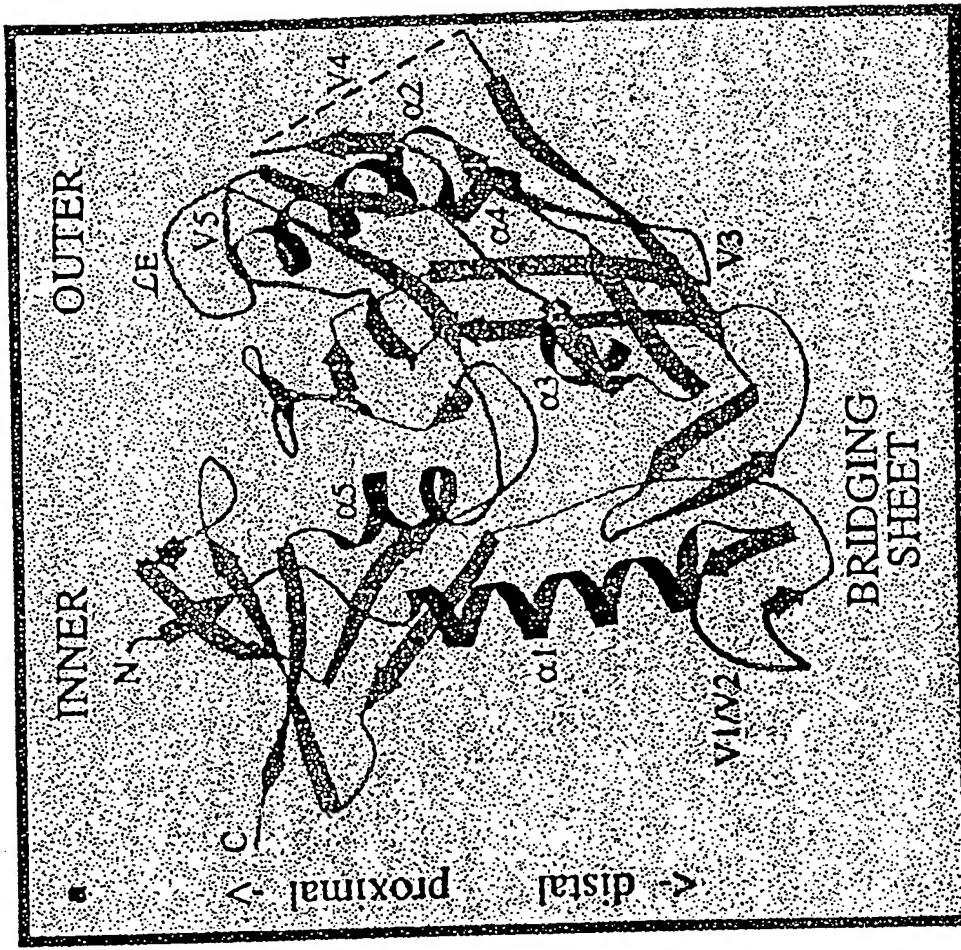


FIG. 1

1 50

HXB2 (1) M~~EVK~~---E~~K~~QHLWRW~~E~~WRWGT~~V~~LL~~G~~LMIC-SATEK~~W~~~~V~~~~G~~VPVWK  
162 (1) ---MDAM~~K~~~~R~~~~L~~CCV~~L~~LC~~G~~SPS~~E~~VEK~~W~~~~V~~~~G~~VPVWK  
SF2 (1) M~~V~~KGTRRN~~E~~QHLWRW~~G~~---T~~L~~L~~G~~LMIC-SATEK~~W~~~~V~~~~G~~VPVWK  
CM236 (1) M~~V~~KETQMN~~V~~PNLW~~K~~W~~G~~---T~~L~~L~~G~~LMIC-S~~S~~NN~~V~~~~V~~~~G~~VPVWK  
US4 (1) --~~MR~~---KHCQHLWRGG~~G~~---I~~L~~L~~G~~LMIC-RE~~T~~TV~~R~~~~V~~~~G~~VPVWK  
Consensus (1) MRVK YQHLWRWG TLLL~~G~~LMIC SATEKLWVTVYYGVPVWK

51 100

HXB2 (47) EATT~~T~~FCASDAKAYDT~~E~~V~~H~~N~~V~~W~~A~~~~T~~~~H~~AC~~V~~PT~~D~~PN~~P~~QE~~V~~V~~L~~ N~~V~~T~~E~~N~~F~~N~~M~~  
162 (41) EATT~~T~~FCASDAKAYDT~~E~~V~~H~~N~~V~~W~~A~~~~T~~~~H~~AC~~V~~PT~~D~~PN~~P~~QE~~V~~V~~L~~ N~~V~~T~~E~~N~~F~~N~~M~~  
SF2 (46) EATT~~T~~FCASDAKAYDT~~E~~V~~H~~N~~V~~W~~A~~~~T~~~~H~~AC~~V~~PT~~D~~PN~~P~~QE~~V~~V~~L~~ N~~V~~T~~E~~N~~F~~N~~M~~  
CM236 (46) DAD~~T~~FCASDAKAYDT~~E~~V~~H~~N~~V~~W~~A~~~~T~~~~H~~AC~~V~~PT~~D~~PN~~P~~QE~~V~~V~~L~~ N~~V~~T~~E~~N~~F~~N~~M~~  
US4 (41) EATT~~T~~FCASDAKAYKABARN~~V~~AC~~V~~PT~~D~~PN~~P~~QE~~V~~V~~L~~ N~~V~~T~~E~~N~~F~~N~~M~~  
Consensus (51) EATT~~T~~FCASDAKAYDT~~E~~V~~H~~N~~V~~W~~A~~~~T~~~~H~~AC~~V~~PT~~D~~PN~~P~~QE~~V~~V~~L~~ N~~V~~T~~E~~N~~F~~N~~M~~

101 150

HXB2 (97) KND~~V~~YQMHEDI~~I~~SLWDQSLKPCV~~K~~L~~T~~PLCV~~T~~LN~~C~~DL~~-----~~  
162 (91) KNNMVEQM~~H~~~~E~~~~D~~~~I~~SLWDQSLKPCV~~K~~L~~T~~PLCV~~T~~LN~~C~~DL~~-----~~  
SF2 (96) KNNMVEQM~~H~~~~E~~~~D~~~~I~~SLWDQSLKPCV~~K~~L~~T~~PLCV~~T~~LN~~C~~NAK~~-----~~  
CM236 (96) KNNMVEQM~~H~~~~E~~~~D~~~~I~~SLWDQSLKPCV~~K~~L~~T~~PLCV~~T~~LN~~C~~NAK~~-----~~  
US4 (91) KNNMVEQM~~H~~~~E~~~~D~~~~I~~SLWDQSLKPCV~~K~~L~~T~~PLCV~~T~~LN~~C~~DKLTGSTNGTNSTS  
Consensus (101) KNNMVEQM~~H~~~~E~~~~D~~~~I~~SLWDQSLKPCV~~K~~L~~T~~PLCV~~T~~LN~~C~~DL

151 200

HXB2 (135) -----KND~~T~~NTN~~E~~SSG~~P~~MI~~E~~KEIK~~N~~EN~~I~~TS~~S~~IR~~G~~Q~~E~~Y~~A~~LF~~Y~~  
162 (129) -----K~~N~~D~~T~~NTK~~E~~SNW~~E~~MD~~-~~KEIK~~N~~EN~~I~~TS~~S~~IR~~G~~Q~~E~~Y~~A~~LF~~Y~~  
SF2 (134) -----GK~~N~~D~~T~~NTN~~E~~SNW~~E~~EE~~-~~KEIK~~N~~EN~~I~~TS~~S~~IR~~G~~Q~~E~~Y~~A~~LF~~Y~~  
CM236 (135) -----LT~~N~~V~~N~~N~~I~~TS~~S~~NT~~G~~IGN~~I~~TD~~V~~W~~N~~S~~N~~T~~E~~DK~~D~~K~~V~~H~~A~~LF~~Y~~  
US4 (141) GT~~N~~STSGT~~N~~EST~~N~~SD~~S~~WEK~~M~~PEG~~E~~IK~~N~~CS~~N~~IT~~S~~SD~~D~~Q~~E~~Y~~A~~LF~~Y~~  
Consensus (151) NAT~~N~~NT~~N~~SS KE M KGEIKNC~~S~~FN~~I~~TT~~S~~IR~~D~~K~~V~~Q~~E~~Y~~A~~LF~~Y~~

201 250

HXB2 (178) K~~D~~V~~V~~D~~N~~DT~~T~~TS~~-----~~KE~~G~~T~~S~~NT~~S~~W~~T~~AK~~E~~K~~V~~S~~D~~E~~-----~~GA~~P~~  
162 (171) K~~D~~V~~V~~D~~N~~NT~~S~~TS~~-----~~KE~~G~~IN~~N~~NT~~S~~W~~T~~AK~~E~~K~~V~~S~~D~~E~~-----~~GA~~P~~  
SF2 (176) N~~D~~V~~V~~D~~N~~AST~~T~~NY~~T~~N~~R~~~~I~~H~~N~~R~~E~~~~S~~T~~-----~~KE~~G~~IN~~N~~NT~~S~~W~~T~~AK~~E~~K~~V~~S~~D~~E~~-----~~GA~~P~~  
CM236 (179) K~~D~~V~~V~~D~~N~~DK~~T~~S~~-----~~SE~~G~~IN~~N~~NT~~S~~W~~T~~AK~~E~~K~~V~~S~~D~~E~~-----~~GA~~P~~  
US4 (191) K~~D~~V~~V~~D~~N~~DN~~N~~AS~~-----~~KE~~G~~IN~~N~~NT~~S~~W~~T~~AK~~E~~K~~V~~S~~D~~E~~-----~~GA~~P~~  
Consensus (201) KLDVV~~V~~PIDND TS YRLINCNTSVITQACPKVS~~F~~EP~~I~~PI~~H~~Y~~C~~AP~~G~~

251 300

HXB2 (223) F~~A~~IL~~K~~NN~~K~~TE~~G~~T~~G~~T~~N~~S~~T~~TC~~G~~U~~G~~TC~~R~~V~~S~~SG~~N~~TS~~S~~AA~~E~~EV~~V~~  
162 (216) F~~A~~IL~~K~~NN~~K~~ND~~K~~NG~~S~~TC~~G~~T~~N~~S~~T~~TC~~G~~U~~G~~TC~~R~~V~~S~~SG~~N~~TS~~S~~AA~~E~~EV~~V~~  
SF2 (226) F~~A~~IL~~K~~NN~~K~~TE~~G~~K~~N~~K~~S~~TC~~G~~T~~N~~S~~T~~TC~~G~~U~~G~~TC~~R~~V~~S~~SG~~N~~TS~~S~~AA~~E~~EV~~V~~  
CM236 (226) F~~A~~IL~~K~~NN~~K~~ND~~K~~NG~~S~~TC~~G~~K~~N~~K~~S~~TC~~G~~T~~N~~S~~T~~TC~~G~~U~~G~~TC~~R~~V~~S~~SG~~N~~TS~~S~~AA~~E~~EV~~V~~  
US4 (236) F~~A~~IL~~K~~ND~~K~~KG~~N~~CT~~G~~K~~N~~K~~S~~TC~~G~~T~~N~~S~~T~~TC~~G~~U~~G~~TC~~R~~V~~S~~SG~~N~~TS~~S~~AA~~E~~EV~~V~~  
Consensus (251) FAILKCNDK FNGTGPCTNVSTVQCTH~~G~~IR~~P~~V~~V~~ST~~Q~~LL~~N~~GS~~A~~EE~~V~~

301 350

HXB2 (273) R~~S~~V~~N~~F~~D~~~~M~~~~K~~~~Q~~~~E~~Q~~N~~T~~E~~V~~E~~W~~G~~T~~R~~N~~N~~K~~R~~R~~I~~ORG~~G~~GRA~~E~~VT~~I~~~~K~~  
162 (266) R~~S~~E~~N~~F~~D~~~~N~~~~K~~~~Q~~~~E~~Q~~N~~T~~E~~V~~E~~W~~G~~T~~R~~N~~N~~K~~R~~S~~T~~~~E~~---~~G~~GRA~~F~~Y~~A~~T~~E~~D  
SF2 (276) R~~S~~T~~N~~F~~D~~~~N~~~~K~~~~Q~~~~E~~Q~~N~~T~~E~~V~~E~~W~~G~~T~~R~~N~~N~~K~~R~~S~~T~~~~E~~---~~G~~GRA~~F~~Y~~A~~T~~E~~R  
CM236 (276) R~~S~~E~~N~~L~~N~~F~~D~~~~N~~~~K~~~~Q~~~~E~~Q~~N~~T~~E~~V~~E~~W~~G~~T~~R~~N~~N~~K~~R~~S~~T~~~~E~~---~~G~~Q~~V~~E~~Y~~R~~T~~D  
US4 (286) R~~S~~E~~N~~F~~D~~~~N~~~~K~~~~Q~~~~E~~Q~~N~~T~~E~~V~~E~~W~~G~~T~~R~~N~~N~~K~~R~~S~~T~~~~E~~---~~G~~GRA~~F~~Y~~A~~T~~E~~D  
Consensus (301) RSENFTDNAK~~T~~IIVQLN~~E~~S~~V~~E~~I~~N~~C~~TR~~P~~NN~~N~~TR~~K~~SI I GPGRAF~~Y~~ TGD

FIG. 2A

351

400

HXB2 (323) I-GNMQAHENISRAKWNNTKQIASKIREQEGNNKTIIEKQSSGGPPEI  
 162 (314) IIEDIQQAHENESGEKWNNTKQIVTQQAQFG-NKRAVEKQSSGGPPEI  
 SF2 (324) IIEDIQQAHENISRAQWNNTKQIVKIREQEGNNKTIIEKQSSGGPPEI  
 CM236 (324) IIEDIQQAHENESGEKWNNTKQIVTQQAQFG-NKRAVEKQSSGGPPEI  
 US4 (334) IIEDIQQAHENESKANNTNTKQIVKIREQEGNNKTIIEKQSSGGPPEI  
 Consensus (351) IIGDIRQAHCNISRAKWNNTL QIV KLREQFGNNKTIIFNQSSGGDPEI

401

450

HXB2 (372) VTIISENGGEPEKNTQDENSWFNSTWSTEGSNNTEGSDITTEGREGK  
 162 (363) VMHSENGGEPEKNTQDENSWF-NN--TIGPNNTNG--NITTEGREGK  
 SF2 (374) VMHSENGGEPEKNTQDENSWFRLN--HTEG--TKGNDITTEGREGK  
 CM236 (373) TMHENGGEPEKNTTRNNCIEN--GEMG--GCNG--NITTEGREGK  
 US4 (384) VTIISENGGEPEKNTQDENSW--N--IIEEVNKTKEKDITTEGREGK  
 Consensus (401) VMHSFNCGGEFFYCNTQLFNSTW N TEG N T G DTIILPCRIK

↓

451

↓

500

HXB2 (422) QIINMWQEVGKAMYAPPI GQIRCSSNITGLLTDGG--NSNNESTI  
 162 (407) QIINMWQEVGKAMYAPPI GQIRCSSNITGLLTDGG--EISNTT  
 SF2 (419) QIINMWQEVGKAMYAPPI GQIRCSSNITGLLTDGG--NNTNT  
 CM236 (417) QIINMWQEVGKAMYAPPI GQIRCSSNITGLLTDGG--AINTNT  
 US4 (430) QIINMWQEVGKAMYAPPI GQIRCSSNITGLLTDGG--NNTNT  
 Consensus (451) QIINMWQEVGKAMYAPPI GQIRCSSNITGLLTDGG NITNDTEIF

501

550

HXB2 (469) RGGGGDMRDNRSELYKVVKIEPLGVAPTKAKRRVVQREKRAVGI--  
 162 (455) RGGGGDMRDNRSELYKVVKIEPLGVAPTKAKRRVVQREKRAVGI--  
 SF2 (467) RGGGGDMRDNRSELYKVVKIEPLGVAPTKAKRRVVQREKRAVGI--  
 CM236 (464) RGGGGDMRDNRSELYKVVKIEPLGVAPTKAKRRVVQREKRAVGI--  
 US4 (480) RGGGGDMRDNRSELYKVVKIEPLGVAPTKAKRRVVQREKRAVGI--  
 Consensus (501) RGGGGDMRDNRSELYKVVKIEPLGVAPTKAKRRVVQREKRAVGI GA

551

600

HXB2 (518) MFLGFLGAAGSTMGAASLTQVQARQLLSGIVQQQNNLLRAIEAQHLLQ  
 162 (504) MFLGFLGAAGSTMGAASLTQVQARQLLSGIVQQQNNLLRAIEAQHLLQ  
 SF2 (517) MFLGFLGAAGSTMGAASLTQVQARQLLSGIVQQQNNLLRAIEAQHLLQ  
 CM236 (513) MFLGFLGAAGSTMGAASLTQVQARQLLSGIVQQQNNLLRAIEAQHLLQ  
 US4 (529) MFLGFLGAAGSTMGAASLTQVQARQLLSGIVQQQNNLLRAIEAQHLLQ  
 Consensus (551) MFLGFLGAAGSTMGAASLTQVQARQLLSGIVQQQNNLLRAIEAQHLLQ

601

650

HXB2 (568) LTVWGIKQLQARVLAVERYLKDQQLGIWCGSGKLICTTAVPNASWSNK  
 162 (554) LTVWGIKQLQARVLAVERYLKDQQLGIWCGSGKLICTTAVPNASWSNK  
 SF2 (567) LTVWGIKQLQARVLAVERYLKDQQLGIWCGSGKLICTTAVPNASWSNK  
 CM236 (563) LTVWGIKQLQARVLAVERYLKDQQLGIWCGSGKLICTTAVPNASWSNK  
 US4 (579) LTVWGIKQLQARVLAVERYLKDQQLGIWCGSGKLICTTAVPNASWSNK  
 Consensus (601) LTVWGIKQLQARVLAVERYLKDQQLGIWCGSGKLICTTAVPNASWSNK

FIG. 2B

651

700

HXB2 (618) SLEQWANHTWMENDRGHNNTYSLIISLIEESQNCQEKNEQE VIGEERDKWA  
 162 (604) SLDQWNNMWMENRERDNYNTLYLIEESQNCQEKNEQE FICEDDKWA  
 SF2 (617) SLEQWANMWMQNE RERDNYNTLYLIEESQNCQEKNEQE IAGEDDKWA  
 CM236 (613) SYEQWNNMWMENRERDNYNTYLYEETESQNCQEKNEQE KQEDDKWA  
 US4 (629) SLTQWANMWMENRERDNYGLIYNLIEIAQNCQEKNEQE LIGEDDKWA  
 Consensus (651) SLEELWNNMWMWEREI NYTNLIYLIEESQNCQEKNEQELLELDKWA

701

750

HXB2 (668) SAWNWFNTNWYKQDPMIVSGIVSLRIVBAQSLWANVRQGSPLSF  
 162 (654) SAWNWFDTKQWYKIKINWYVSGIVSLRIVBAQSLWANVRQGSPLSF  
 SF2 (667) SAWNWFDTKQWYKIKI DPMIVSGIVSLRIVBAQSLWANVRQGSPLSF  
 CM236 (663) SAWNWFDTKQWYKIKI DPMIVSGIVSLRIVBAQSLWANVRQGSPLSF  
 US4 (679) SAWNWFDTNQWYKIKI DPMIVSGIVSLRIVBAQSLWANVRQGSPLSF  
 Consensus (701) SLWNWFDTNWLYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGSPLSF

751

•800

HXB2 (718) QHLPTPQGPDPEGEEEGGERDQDRSIVNSFQIADDERSNS  
 162 (704) QRLPAPQGPDPEGEEEGGERDQDRSSPQHLSLQIADDERSNS  
 SF2 (717) QRLPVPQGPDPEGEEEGGERDQDRSIVNSFQIADDERSNS  
 CM236 (713) QPFHHQPEPDSERQGQGDRSIVNSFQIADDERSNS  
 US4 (729) QRLPAQGPDPEGEEEGGERDQDRSIVNSFQIADDERSNS  
 Consensus (751) QTRLP PRGPDRPEGIEEGGERDQDRSIVNSFQIADDERSNS

801

850

HXB2 (768) YHRLDPLLAVTRIVELGR-----RGWEALKYWWNLLQYQELKNS  
 162 (754) YHRLDPLLAVIAAIVELGR-----RGWEALKYWWNLLQYQELKNS  
 SF2 (767) YHRLDPLLAVIAARTVIELGR-----RGWEALKYWWNLLQYQELKNS  
 CM236 (763) YHRLDPLLAVIAARTVIELGR-----RGWEALKYWWNLLQYQELKNS  
 US4 (779) YHRLDPLLAVARIVELGR-----RGWEALKYWWNLLQYQELKNS  
 Consensus (801) YHRLDPLLAVARIVELLGR RGWEALKYWWNLLQYQELKNS

851

900

HXB2 (811) AVSLNATAIAAEGTDRVIEVAQRAFRAILHIPRRIQGLER LL  
 162 (797) AVLFDTIAIAAEGTDRVIEVAFRAILIPRRIQGLER LL  
 SF2 (810) AVWLNTIAIAAEGTDRVIEVAFRAILIPRRIQGLER LL  
 CM236 (813) AVLLDTIAIAAEGTDRVIEVAQGAWRILIPRRIQGLER LL  
 US4 (822) AVLFNTIAIAAEGTDRVIEVAQRAFRAILHIPRRIQGLER LL  
 Consensus (851) AVSLNATAIAAEGTDRVIEVAQRAFRAILHIPRRIQGLER LL

FIG. 2C

	1	40
Leu122-Ser199	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val127-Asn195	(1)	GAATTGCCACCATGGATCCAATGAAGAGAGGGCTCTGCT
Val120-Ile201B	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Ala204	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Ile201	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Thr202	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Lys121-Val200	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Consensus	(1)	GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
	41	80
Leu122-Ser199	(41)	GTGTGCTGCTGGTGTGGAGCAGTCTCGTTTCGCCAG
Val127-Asn195	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
Val120-Ile201B	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
Val120-Ala204	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
Val120-Ile201	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
Val120-Thr202	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
Lys121-Val200	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
Consensus	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTCGTTTCGCCAG
	81	120
Leu122-Ser199	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val127-Asn195	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201B	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ala204	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Thr202	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Lys121-Val200	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Consensus	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
	121	160
Leu122-Ser199	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Val127-Asn195	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Val120-Ile201B	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Val120-Ala204	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Val120-Ile201	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Val120-Thr202	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Lys121-Val200	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
Consensus	(121)	CCCGTGTGGAGGAGGCCACCAACCCCTGTTCTGCCCA
	161	200
Leu122-Ser199	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Val127-Asn195	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Val120-Ile201B	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Val120-Ala204	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Val120-Ile201	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Val120-Thr202	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Lys121-Val200	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
Consensus	(161)	GGGACGCCAGGCTTACGGACACCGAGGTGCACAACTGTGTG
	201	240
Leu122-Ser199	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Val127-Asn195	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Val120-Ile201B	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Val120-Ala204	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Val120-Ile201	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Val120-Thr202	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Lys121-Val200	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
Consensus	(201)	GGCACCCACGCCCTGGCTCCCCACGGACCCCAACCCCCAG
	241	280
Leu122-Ser199	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
Val127-Asn195	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT

Val120-Ile2018	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
Val120-Ala204	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
Val120-Ile201	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
Val120-Thr202	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
Lys121-Val200	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
Consensus	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACTTCAACATGT
	281	320
Leu122-Ser199	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Val127-Asn195	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Val120-Ile2018	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Val120-Ala204	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Val120-Ile201	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Val120-Thr202	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Lys121-Val200	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
Consensus	(281)	GGAAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCAT
	321	360
Leu122-Ser199	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Val127-Asn195	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Val120-Ile2018	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCCTGCC---
Val120-Ala204	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGG---
Val120-Ile201	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGG---
Val120-Thr202	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGG---
Lys121-Val200	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGG-
Consensus	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGG-
	361	400
Leu122-Ser199	(361)	-----GGCAA-----CAGCG
Val127-Asn195	(361)	ACCCCCCTGTGCGTGGGGCAGGGAACTGCAACACCAAGCG
Val120-Ile2018	(357)	-----CG
Val120-Ala204	(357)	-----
Val120-Ile201	(357)	-----CG
Val120-Thr202	(357)	-----CG
Lys121-Val200	(359)	-----C-----CCCCG
Consensus	(361)	-----CG
	401	440
Leu122-Ser199	(371)	TGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Val127-Asn195	(401)	TGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Val120-Ile2018	(359)	GCATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Val120-Ala204	(357)	-----CGCCGGCCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Val120-Ile201	(359)	GCATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Val120-Thr202	(359)	GCGCCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Lys121-Val200	(365)	TGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
Consensus	(401)	ATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCAT
	441	480
Leu122-Ser199	(411)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Val127-Asn195	(441)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Val120-Ile2018	(399)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Val120-Ala204	(393)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Val120-Ile201	(399)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Val120-Thr202	(399)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Lys121-Val200	(405)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
Consensus	(441)	CCCCATCCACTACTGGCCCCCCCCGGGCTTCGCCATCCGTG
	481	520
Leu122-Ser199	(451)	AAAGTCAACGACAAGAAGTCAACGGCAGCGGCCCTGCA
Val127-Asn195	(481)	AAAGTCAACGACAAGAAGTCAACGGCAGCGGCCCTGCA
Val120-Ile2018	(439)	AAAGTCAACGACAAGAAGTCAACGGCAGCGGCCCTGCA
Val120-Ala204	(433)	AAAGTCAACGACAAGAAGTCAACGGCAGCGGCCCTGCA
Val120-Ile201	(439)	AAAGTCAACGACAAGAAGTCAACGGCAGCGGCCCTGCA

Val120-Thr202	(439)	AAGTGCACGACAAGAAGTCAACGGCAGCGGCCCTGCA
Lys121-Val1200	(445)	AAGTGCACGACAAGAAGTCAACGGCAGCGGCCCTGCA
Consensus	(481)	AAGTGCACGACAAGAAGTCAACGGCAGCGGCCCTGCA
		560
Leu122-Ser199	(491)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Val127-Asn195	(521)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Val120-Ile201B	(479)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Val120-Ala204	(473)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Val120-Ile201	(479)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Val120-Thr202	(479)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Lys121-Val1200	(485)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
Consensus	(521)	CCAAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC
		600
Leu122-Ser199	(531)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Val127-Asn195	(561)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201B	(519)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ala204	(513)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201	(519)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Thr202	(519)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Lys121-Val1200	(525)	CGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCC
Consensus	(561)	CGTGGTGAGCACCAGCTGCTGAACGGCAGCCTGGCC
		640
Leu122-Ser199	(571)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Val127-Asn195	(601)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Ile201B	(559)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Ala204	(553)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Ile201	(559)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Thr202	(559)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Lys121-Val1200	(565)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
Consensus	(601)	GAGGAGGGCTGGTGATCCGCAGCGAGAACTTCACCGACA
		680
Leu122-Ser199	(611)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Val127-Asn195	(641)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Val120-Ile201B	(599)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Val120-Ala204	(593)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Val120-Ile201	(599)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Val120-Thr202	(599)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Lys121-Val1200	(605)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
Consensus	(641)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGG
		720
Leu122-Ser199	(651)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Val127-Asn195	(681)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Val120-Ile201B	(639)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Val120-Ala204	(633)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Val120-Ile201	(639)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Val120-Thr202	(639)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Lys121-Val1200	(645)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
Consensus	(681)	GATCAACTGCACCCGCCCAAGAACAAACACCCGCAAGAGC
		760
Leu122-Ser199	(691)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Val127-Asn195	(721)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Val120-Ile201B	(679)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Val120-Ala204	(673)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Val120-Ile201	(679)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Val120-Thr202	(679)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Lys121-Val1200	(685)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG
Consensus	(721)	ATCACCATCGGCCCCGGCGCGCCTCTACGCACCGGGCG

		761	800
Leu122-Ser199	(731)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Val127-Asn195	(761)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Val120-Ile201B	(719)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Val120-Ala204	(713)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Val120-Ile201	(719)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Val120-Thr202	(719)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Lys121-Val1200	(725)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
Consensus	(761)	ACATCATGGCGACATCCGCAGGCCACTGCAACATCAG	
		801	840
Leu122-Ser199	(771)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Val127-Asn195	(801)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Val120-Ile201B	(759)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Val120-Ala204	(753)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Val120-Ile201	(759)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Val120-Thr202	(759)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Lys121-Val1200	(765)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
Consensus	(801)	CGCGAGAAGTGGAACAAACCCCTGAAGCAGATCGTGACC	
		841	880
Leu122-Ser199	(811)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Val127-Asn195	(841)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Val120-Ile201B	(799)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Val120-Ala204	(793)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Val120-Ile201	(799)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Val120-Thr202	(799)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Lys121-Val1200	(805)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
Consensus	(841)	AAGCTGCAGGCCAGTTGGCAACAAGACCATCGTGTCA	
		881	920
Leu122-Ser199	(851)	AGCAGAGCAGCGGGCGAGCCCGAGATCGTGTACAG	
Val127-Asn195	(881)	AGCAGAGCAGCGGGCGAGCCCGAGATCGTGTACAG	
Val120-Ile201B	(839)	AGCAGAGCAGCGGGGGCGACCCCGAGATCGTGTACAG	
Val120-Ala204	(833)	AGCAGAGCAGCGGGGGCGACCCCGAGATCGTGTACAG	
Val120-Ile201	(839)	AGCAGAGCAGCGGGGGCGACCCCGAGATCGTGTACAG	
Val120-Thr202	(839)	AGCAGAGCAGCGGGGGCGACCCCGAGATCGTGTACAG	
Lys121-Val1200	(845)	AGCAGAGCAGCGGGGGCGACCCCGAGATCGTGTACAG	
Consensus	(881)	AGCAGAGCAGCGGGGGCGACCCCGAGATCGTGTACAG	
		921	960
Leu122-Ser199	(891)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Val127-Asn195	(921)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Val120-Ile201B	(879)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Val120-Ala204	(873)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Val120-Ile201	(879)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Val120-Thr202	(879)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Lys121-Val1200	(885)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
Consensus	(921)	CTTCAACTGGCGCGCGAGTTCTTACTGCAACAGCACC	
		961	1000
Leu122-Ser199	(931)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Val127-Asn195	(961)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Val120-Ile201B	(919)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Val120-Ala204	(913)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Val120-Ile201	(919)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Val120-Thr202	(919)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Lys121-Val1200	(925)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
Consensus	(961)	CAGCTGTTCAACAGCACCTGGAAACACACCATCGGCCCCA	
		1001	1040
Leu122-Ser199	(971)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA	
Val127-Asn195	(1001)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA	

Val120-Ile201B	(959)	ACAACACCAACGGCACCACATCACCTGCCCTGCCGCATCAA
Val120-Ala204	(953)	ACAACACCAACGGCACCACATCACCTGCCCTGCCGCATCAA
Val120-Ile201	(959)	ACAACACCAACGGCACCACATCACCTGCCCTGCCGCATCAA
Val120-Thr202	(959)	ACAACACCAACGGCACCACATCACCTGCCCTGCCGCATCAA
Lys121-Val200	(965)	ACAACACCAACGGCACCACATCACCTGCCCTGCCGCATCAA
Consensus	(1001)	ACAACACCAACGGCACCACATCACCTGCCCTGCCGCATCAA
	1041	1080
Leu122-Ser199	(1011)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val127-Asn195	(1041)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ile201B	(999)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ala204	(993)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ile201	(999)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Thr202	(999)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Lys121-Val200	(1005)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Consensus	(1041)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
	1081	1120
Leu122-Ser199	(1051)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Val127-Asn195	(1081)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201B	(1039)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Val120-Ala204	(1033)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201	(1039)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Val120-Thr202	(1039)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Lys121-Val200	(1045)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
Consensus	(1081)	TACGCCCCCCCCCATCCGGGCCAGATCCGCTGCAGCAGCA
	1121	1160
Leu122-Ser199	(1091)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Val127-Asn195	(1121)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Val120-Ile201B	(1079)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Val120-Ala204	(1073)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Val120-Ile201	(1079)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Val120-Thr202	(1079)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Lys121-Val200	(1085)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
Consensus	(1121)	ACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGA
	1161	1200
Leu122-Ser199	(1131)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Val127-Asn195	(1161)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Val120-Ile201B	(1119)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Val120-Ala204	(1113)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Val120-Ile201	(1119)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Val120-Thr202	(1119)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Lys121-Val200	(1125)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
Consensus	(1161)	GATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGC
	1201	1240
Leu122-Ser199	(1171)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Val127-Asn195	(1201)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201B	(1159)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ala204	(1153)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201	(1159)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Thr202	(1159)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Lys121-Val200	(1165)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
Consensus	(1201)	GACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACA
	1241	1280
Leu122-Ser199	(1211)	AGGTGGTGAAGATCGAGCCCCCTGGCGTGGCCCCCACCAA
Val127-Asn195	(1241)	AGGTGGTGAAGATCGAGCCCCCTGGCGTGGCCCCCACCAA
Val120-Ile201B	(1199)	AGGTGGTGAAGATCGAGCCCCCTGGCGTGGCCCCCACCAA
Val120-Ala204	(1193)	AGGTGGTGAAGATCGAGCCCCCTGGCGTGGCCCCCACCAA
Val120-Ile201	(1199)	AGGTGGTGAAGATCGAGCCCCCTGGCGTGGCCCCCACCAA

Val120-Thr202	(1199)	AGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCACAA	
Lys121-Val200	(1205)	AGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCACAA	
Consensus	(1241)	AGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCACAA	
		1281	1320
Leu122-Ser199	(1251)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Val127-Asn195	(1281)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Val120-Ile201B	(1239)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Val120-Ala204	(1233)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Val120-Ile201	(1239)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Val120-Thr202	(1239)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Lys121-Val200	(1245)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
Consensus	(1281)	GGCCAAGCGCCCGCGTGGTGCAGCGCGAGAACCGCGCCGTG	
		1321	1360
Leu122-Ser199	(1291)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Val127-Asn195	(1321)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Val120-Ile201B	(1279)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Val120-Ala204	(1273)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Val120-Ile201	(1279)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Val120-Thr202	(1279)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Lys121-Val200	(1285)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
Consensus	(1321)	ACCTGGCGCCATGTTCTGGCCTTCCTGGCGCCGCCG	
		1361	1400
Leu122-Ser199	(1331)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Val127-Asn195	(1361)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Val120-Ile201B	(1319)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Val120-Ala204	(1313)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Val120-Ile201	(1319)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Val120-Thr202	(1319)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Lys121-Val200	(1325)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
Consensus	(1361)	GCAGCACCATGGCGCCCGCAGCCCTGACCCCTGACCGTGC	
		1401	1440
Leu122-Ser199	(1371)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Val127-Asn195	(1401)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Val120-Ile201B	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Val120-Ala204	(1353)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Val120-Ile201	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Val120-Thr202	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Lys121-Val200	(1365)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
Consensus	(1401)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCACCAAC	
		1441	1480
Leu122-Ser199	(1411)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Val127-Asn195	(1441)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Val120-Ile201B	(1399)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Val120-Ala204	(1393)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Val120-Ile201	(1399)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Val120-Thr202	(1399)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Lys121-Val200	(1405)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
Consensus	(1441)	AACTGCTGGCGCCATCGAGGCCAGCAGCACCTGCTGC	
		1481	1520
Leu122-Ser199	(1451)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Val127-Asn195	(1481)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Val120-Ile201B	(1439)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Val120-Ala204	(1433)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Val120-Ile201	(1439)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Val120-Thr202	(1439)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Lys121-Val200	(1445)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	
Consensus	(1481)	AGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCT	

		1521	1560
Leu122-Ser199	(1491)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Val127-Asn195	(1521)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Val120-Ile201B	(1479)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Val120-Ala204	(1473)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Val120-Ile201	(1479)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Val120-Thr202	(1479)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Lys121-Val200	(1485)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
Consensus	(1521)	GCTGCCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTG	
		1561	1600
Leu122-Ser199	(1531)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Val127-Asn195	(1561)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Val120-Ile201B	(1519)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Val120-Ala204	(1513)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Val120-Ile201	(1519)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Val120-Thr202	(1519)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Lys121-Val200	(1525)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
Consensus	(1561)	GGCATCTGGGGCTGCAGCGGAAAGCTGATCTGCACCAACCG	
		1601	1640
Leu122-Ser199	(1571)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Val127-Asn195	(1601)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Val120-Ile201B	(1559)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Val120-Ala204	(1553)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Val120-Ile201	(1559)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Val120-Thr202	(1559)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Lys121-Val200	(1565)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
Consensus	(1601)	CCGTGCCCTGAAACGCCAGCTGGAGCAACAAGAGCCTGGA	
		1641	1680
Leu122-Ser199	(1611)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Val127-Asn195	(1641)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Val120-Ile201B	(1599)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Val120-Ala204	(1593)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Val120-Ile201	(1599)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Val120-Thr202	(1599)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Lys121-Val200	(1605)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
Consensus	(1641)	CCAGATCTGAAACAACATGACCTGGATGGAGTGGAGCTGGAGC	
		1681	1720
Leu122-Ser199	(1651)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Val127-Asn195	(1681)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Val120-Ile201B	(1639)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Val120-Ala204	(1633)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Val120-Ile201	(1639)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Val120-Thr202	(1639)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Lys121-Val200	(1645)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
Consensus	(1681)	GAGATCGACAACCTACACCAACCTGATCTACACCCCTGATCG	
		1721	1760
Leu122-Ser199	(1691)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Val127-Asn195	(1721)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Val120-Ile201B	(1679)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Val120-Ala204	(1673)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Val120-Ile201	(1679)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Val120-Thr202	(1679)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Lys121-Val200	(1685)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
Consensus	(1721)	AGGAGAGCCAGAACCGAGCAGGAGAAGAACGAGCAGGAGCT	
		1761	1800
Leu122-Ser199	(1731)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC	
Val127-Asn195	(1761)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC	

Val120-Ile2018	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC
Val120-Ala204	(1713)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC
Val120-Ile201	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC
Val120-Thr202	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC
Lys121-Val200	(1725)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC
Consensus	(1761)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACGGTTC
	1801	1840
Leu122-Ser199	(1771)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val127-Asn195	(1801)	GACATCAGGAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ile201B	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ala204	(1753)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ile201	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Thr202	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Lys121-Val200	(1765)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Consensus	(1801)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
	1841	1880
Leu122-Ser199	(1811)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Val127-Asn195	(1841)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Val120-Ile201B	(1799)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Val120-Ala204	(1793)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Val120-Ile201	(1799)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Val120-Thr202	(1799)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Lys121-Val200	(1805)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
Consensus	(1841)	TGATCGTGGGGGGCCTGGTGGGCCATCGTGTTCAC
	1881	1920
Leu122-Ser199	(1851)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Val127-Asn195	(1881)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Val120-Ile201B	(1839)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Val120-Ala204	(1833)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Val120-Ile201	(1839)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Val120-Thr202	(1839)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Lys121-Val200	(1845)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
Consensus	(1881)	CGTGTGAGCATCGTGAACCGCGTSCGCCAGGGCTACAGC
	1921	1960
Leu122-Ser199	(1891)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Val127-Asn195	(1921)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Val120-Ile201B	(1879)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Val120-Ala204	(1873)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Val120-Ile201	(1879)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Val120-Thr202	(1879)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Lys121-Val200	(1885)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
Consensus	(1921)	CCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGGGGC
	1961	2000
Leu122-Ser199	(1931)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Val127-Asn195	(1961)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Val120-Ile201B	(1919)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Val120-Ala204	(1913)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Val120-Ile201	(1919)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Val120-Thr202	(1919)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Lys121-Val200	(1925)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
Consensus	(1961)	CCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGGAGCG
	2001	2040
Leu122-Ser199	(1971)	CGACCCCGACCGCAGCAGCCCCCTGGTGCACGGCTGCTG
Val127-Asn195	(2001)	CGACCCCGACCGCAGCAGCCCCCTGGTGCACGGCTGCTG
Val120-Ile201B	(1959)	CGACCCCGACCGCAGCAGCCCCCTGGTGCACGGCTGCTG
Val120-Ala204	(1953)	CGACCCCGACCGCAGCAGCCCCCTGGTGCACGGCTGCTG
Val120-Ile201	(1959)	CGACCCCGACCGCAGCAGCCCCCTGGTGCACGGCTGCTG

Val120-Thr202	(1959)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Lys121-Val200	(1965)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Consensus	(2001)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
	2041	2080
Leu122-Ser199	(2011)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val127-Asn195	(2041)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Ile201B	(1999)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Ala204	(1993)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Ile201	(1999)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Thr202	(1999)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Lys121-Val200	(2005)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Consensus	(2041)	GCCGTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
	2081	2120
Leu122-Ser199	(2051)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Val127-Asn195	(2081)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Val120-Ile201B	(2039)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Val120-Ala204	(2033)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Val120-Ile201	(2039)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Val120-Thr202	(2039)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Lys121-Val200	(2045)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
Consensus	(2081)	GCTACCACCGCCTGGCGACCTGATCCTGATGCCGCCCG
	2121	2160
Leu122-Ser199	(2091)	CATCGTGGAGCTGCTGGGGCGCCGGGGCTGGGAGGCCCTG
Val127-Asn195	(2121)	CATCGTGGAGCTGCTGGGGCGCCGGGGCTGGGAGGCCCTG
Val120-Ile201B	(2079)	CATCGTGGAGCTGCTGGGGCGCCGGGGCTGGGAGGCCCTG
Val120-Ala204	(2073)	CATCGTGGAGCTGCTGGGGCGCCGGGGCTGGGAGGCCCTG
Val120-Ile201	(2079)	CATCGTGGAGCTGCTGGGGCGCCGGGGCTGGGAGGCCCTG
Val120-Thr202	(2079)	CATCGTGGAGCTGCTGGGGGGGGGGGGCTGGGAGGCCCTG
Lys121-Val200	(2085)	CATCGTGGAGCTGCTGGGGCGCCGGCTGGGAGGCCCTG
Consensus	(2121)	CATCGTGGAGCTGCTGGGGCGCCGGCTGGGAGGCCCTG
	2161	2200
Leu122-Ser199	(2131)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Val127-Asn195	(2161)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Val120-Ile201B	(2119)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Val120-Ala204	(2113)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Val120-Ile201	(2119)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Val120-Thr202	(2119)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Lys121-Val200	(2125)	AAGTACTGGGCCAACCTGCTGGACTGACTGGATCCAGGAGC
Consensus	(2161)	AAGTACTGGGCCAACCTGCTGCAGTACTGGATCCAGGAGC
	2201	2240
Leu122-Ser199	(2171)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Val127-Asn195	(2201)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Val120-Ile201B	(2159)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Val120-Ala204	(2153)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Val120-Ile201	(2159)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Val120-Thr202	(2159)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Lys121-Val200	(2165)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
Consensus	(2201)	TGAAGAACAGGCCCGTGACCTGTGGACGCCATGCCAT
	2241	2280
Leu122-Ser199	(2211)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val127-Asn195	(2241)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Ile201B	(2199)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Ala204	(2193)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Ile201	(2199)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Thr202	(2199)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Lys121-Val200	(2205)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Consensus	(2241)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC

		2281	2320
Leu122-Ser199	(2251)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
Val127-Asn195	(2281)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
Val120-Ile201B	(2239)	CAGCGCATCGGGGGGGCCTTCCTGCACATCCCCCGCCGCA	
Val120-Ala204	(2233)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
Val120-Ile201	(2239)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
Val120-Thr202	(2239)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
Lys121-Val200	(2245)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
Consensus	(2281)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA	
		2321	2360
Leu122-Ser199	(2291)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAGCG	
Val127-Asn195	(2321)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAG--	
Val120-Ile201B	(2279)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAGCG	
Val120-Ala204	(2273)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAG--	
Val120-Ile201	(2279)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAG--	
Val120-Thr202	(2279)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAG--	
Lys121-Val200	(2285)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAGCG	
Consensus	(2321)	TCCGCCAGGGCTTCGAGCCCGGCCCTGCTGTAACTCGAG	
		2361	
Leu122-Ser199	(2331)	TGCT	
Val127-Asn195	(2359)	----	
Val120-Ile201B	(2319)	TGCT	
Val120-Ala204	(2311)	----	
Val120-Ile201	(2317)	----	
Val120-Thr202	(2317)	----	
Lys121-Val200	(2325)	TGCT	
Consensus	(2361)		

FIG. 4A

Trp427-Gly431	(201)	[REDACTED]
Gln422-Tyr435B	(201)	[REDACTED]
Arg426-Gly431	(201)	[REDACTED]
Ile423-Met434	(201)	[REDACTED]
Gln422-Tyr435	(201)	[REDACTED]
Arg426-Lys432	(201)	[REDACTED]
Arg426-Gly431B	(201)	[REDACTED]
Asn425-Lys432	(201)	[REDACTED]
Consensus	(201)	GGCCACCCACGCCCTGCGTCCCCACCGACCCCAACCCCCAG
	241	280
Ile424-Ala433	(241)	[REDACTED]
Trp427-Gly431	(241)	[REDACTED]
Gln422-Tyr435B	(241)	[REDACTED]
Arg426-Gly431	(241)	[REDACTED]
Ile423-Met434	(241)	[REDACTED]
Gln422-Tyr435	(241)	[REDACTED]
Arg426-Lys432	(241)	[REDACTED]
Arg426-Gly431B	(241)	[REDACTED]
Asn425-Lys432	(241)	[REDACTED]
Consensus	(241)	GAGATCGTCTGGAGAACGTGACCGAGAACTTAAACATGT
	281	320
Ile424-Ala433	(281)	[REDACTED]
Trp427-Gly431	(281)	[REDACTED]
Gln422-Tyr435B	(281)	[REDACTED]
Arg426-Gly431	(281)	[REDACTED]
Ile423-Met434	(281)	[REDACTED]
Gln422-Tyr435	(281)	[REDACTED]
Arg426-Lys432	(281)	[REDACTED]
Arg426-Gly431B	(281)	[REDACTED]
Asn425-Lys432	(281)	[REDACTED]
Consensus	(281)	GGAAGAACAAACATGGTGGAGCAGATGCACGAGGGACATCAT
	321	360
Ile424-Ala433	(321)	[REDACTED]
Trp427-Gly431	(321)	[REDACTED]
Gln422-Tyr435B	(321)	[REDACTED]
Arg426-Gly431	(321)	[REDACTED]
Ile423-Met434	(321)	[REDACTED]
Gln422-Tyr435	(321)	[REDACTED]
Arg426-Lys432	(321)	[REDACTED]
Arg426-Gly431B	(321)	[REDACTED]
Asn425-Lys432	(321)	[REDACTED]
Consensus	(321)	CAGCCTGTGGACCAGAGCCTGAAGCCCTGCGTGAAGCTG
	361	400
Ile424-Ala433	(361)	[REDACTED]
Trp427-Gly431	(361)	[REDACTED]
Gln422-Tyr435B	(361)	[REDACTED]
Arg426-Gly431	(361)	[REDACTED]
Ile423-Met434	(361)	[REDACTED]
Gln422-Tyr435	(361)	[REDACTED]
Arg426-Lys432	(361)	[REDACTED]
Arg426-Gly431B	(361)	[REDACTED]
Asn425-Lys432	(361)	[REDACTED]
Consensus	(361)	ACCCCCCTGTGCGTGACCCCTGCACTGCACCAACCTGAAGA
	401	440
Ile424-Ala433	(401)	[REDACTED]
Trp427-Gly431	(401)	[REDACTED]
Gln422-Tyr435B	(401)	[REDACTED]

FIG. 4B

Arg426-Gly431	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(401)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
	441	480
Ile424-Ala433	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(441)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(441)	CCGGGGGAGATCAAGAACTGCAGCTTCAAGGTGACCACC
	481	520
Ile424-Ala433	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(481)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(481)	AGCATCCGAAACAAGATGCAGAAGGAGTACGCCCTGTTCT
	521	560
Ile424-Ala433	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(521)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(521)	ACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAAG
	561	600
Ile424-Ala433	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(561)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(561)	CTACAAGCTGATCAACTGCACACCAAGCGTGATCACCCAG
	601	640
Ile424-Ala433	(601)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(601)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(601)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(601)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(601)	ACCCGCAACACCAAGAGCAGCAACTGGAAGGAGATGGA

FIG. 4C

Gln422-Tyr435	(601)	[REDACTED]
Arg426-Lys432	(601)	[REDACTED]
Arg426-Gly431B	(601)	[REDACTED]
Asn425-Lys432	(601)	[REDACTED]
Consensus	(601)	GCCTGCCCAAGGTGAGCTCGAGCCATCCCCATCCACT
	641	680
Ile424-Ala433	(641)	[REDACTED]
Trp427-Gly431	(641)	[REDACTED]
Gln422-Tyr435B	(641)	[REDACTED]
Arg426-Gly431	(641)	[REDACTED]
Ile423-Met434	(641)	[REDACTED]
Gln422-Tyr435	(641)	[REDACTED]
Arg426-Lys432	(641)	[REDACTED]
Arg426-Gly431B	(641)	[REDACTED]
Asn425-Lys432	(641)	[REDACTED]
Consensus	(641)	ACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGA
	681	720
Ile424-Ala433	(681)	[REDACTED]
Trp427-Gly431	(681)	[REDACTED]
Gln422-Tyr435B	(681)	[REDACTED]
Arg426-Gly431	(681)	[REDACTED]
Ile423-Met434	(681)	[REDACTED]
Gln422-Tyr435	(681)	[REDACTED]
Arg426-Lys432	(681)	[REDACTED]
Arg426-Gly431B	(681)	[REDACTED]
Asn425-Lys432	(681)	[REDACTED]
Consensus	(681)	CAAGAAGTTAACGGCAGCGGCCCCCTGCACCAACGTGAGC
	721	760
Ile424-Ala433	(721)	[REDACTED]
Trp427-Gly431	(721)	[REDACTED]
Gln422-Tyr435B	(721)	[REDACTED]
Arg426-Gly431	(721)	[REDACTED]
Ile423-Met434	(721)	[REDACTED]
Gln422-Tyr435	(721)	[REDACTED]
Arg426-Lys432	(721)	[REDACTED]
Arg426-Gly431B	(721)	[REDACTED]
Asn425-Lys432	(721)	[REDACTED]
Consensus	(721)	ACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCA
	761	800
Ile424-Ala433	(761)	[REDACTED]
Trp427-Gly431	(761)	[REDACTED]
Gln422-Tyr435B	(761)	[REDACTED]
Arg426-Gly431	(761)	[REDACTED]
Ile423-Met434	(761)	[REDACTED]
Gln422-Tyr435	(761)	[REDACTED]
Arg426-Lys432	(761)	[REDACTED]
Arg426-Gly431B	(761)	[REDACTED]
Asn425-Lys432	(761)	[REDACTED]
Consensus	(761)	CCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGGAGGGCGT
	801	840
Ile424-Ala433	(801)	[REDACTED]
Trp427-Gly431	(801)	[REDACTED]
Gln422-Tyr435B	(801)	[REDACTED]
Arg426-Gly431	(801)	[REDACTED]
Ile423-Met434	(801)	[REDACTED]
Gln422-Tyr435	(801)	[REDACTED]
Arg426-Lys432	(801)	[REDACTED]

FIG. 4D

Arg426-Gly431B	(801)	GGTGATCCGAGCGAGAACTTCACCGACAACGCCAAGACC	880
Asn425-Lys432	(801)		
Consensus	(801)		
Ile424-Ala433	(841)		
Trp427-Gly431	(841)		
Gln422-Tyr435B	(841)		
Arg426-Gly431	(841)		
Ile423-Met434	(841)		
Gln422-Tyr435	(841)		
Arg426-Lys432	(841)		
Arg426-Gly431B	(841)		
Asn425-Lys432	(841)		
Consensus	(841)	ATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCA	920
Ile424-Ala433	(881)		
Trp427-Gly431	(881)		
Gln422-Tyr435B	(881)		
Arg426-Gly431	(881)		
Ile423-Met434	(881)		
Gln422-Tyr435	(881)		
Arg426-Lys432	(881)		
Arg426-Gly431B	(881)		
Asn425-Lys432	(881)		
Consensus	(881)	CCCGCCCCAACAAACAACACCCGCAAGAGCATCACCATCGG	960
Ile424-Ala433	(921)		
Trp427-Gly431	(921)		
Gln422-Tyr435B	(921)		
Arg426-Gly431	(921)		
Ile423-Met434	(921)		
Gln422-Tyr435	(921)		
Arg426-Lys432	(921)		
Arg426-Gly431B	(921)		
Asn425-Lys432	(921)		
Consensus	(921)	CCCCGGCCGCGCCTTCTACGCCACCGGCGACATCATCGGC	1000
Ile424-Ala433	(961)		
Trp427-Gly431	(961)		
Gln422-Tyr435B	(961)		
Arg426-Gly431	(961)		
Ile423-Met434	(961)		
Gln422-Tyr435	(961)		
Arg426-Lys432	(961)		
Arg426-Gly431B	(961)		
Asn425-Lys432	(961)		
Consensus	(961)	GACATCCGCCAGGCCACTGCAACATCAGCGGGAGAAGT	1040
Ile424-Ala433	(1001)		
Trp427-Gly431	(1001)		
Gln422-Tyr435B	(1001)		
Arg426-Gly431	(1001)		
Ile423-Met434	(1001)		
Gln422-Tyr435	(1001)		
Arg426-Lys432	(1001)		
Arg426-Gly431B	(1001)		
Asn425-Lys432	(1001)		

FIG. 4E

Consensus	(1001)	GGAACAAACACCCCTGAAGCAGATCGTGACCAAGCTGCAGGC	
	1041		1080
Ile424-Ala433	(1041)	[REDACTED]	
Trp427-Gly431	(1041)	[REDACTED]	
Gln422-Tyr435B	(1041)	[REDACTED]	
Arg426-Gly431	(1041)	[REDACTED]	
Ile423-Met434	(1041)	[REDACTED]	
Gln422-Tyr435	(1041)	[REDACTED]	
Arg426-Lys432	(1041)	[REDACTED]	
Arg426-Gly431B	(1041)	[REDACTED]	
Asn425-Lys432	(1041)	[REDACTED]	
Consensus	(1041)	CCAGTCGGCAACAAGACCATCGTGTCAAGCAGAGCAGC	
	1081		1120
Ile424-Ala433	(1081)	[REDACTED]	
Trp427-Gly431	(1081)	[REDACTED]	
Gln422-Tyr435B	(1081)	[REDACTED]	
Arg426-Gly431	(1081)	[REDACTED]	
Ile423-Met434	(1081)	[REDACTED]	
Gln422-Tyr435	(1081)	[REDACTED]	
Arg426-Lys432	(1081)	[REDACTED]	
Arg426-Gly431B	(1081)	[REDACTED]	
Asn425-Lys432	(1081)	[REDACTED]	
Consensus	(1081)	GGCGGCGACCCCGAGATCGTGTGCACAGCTTCAACTGCG	
	1121		1160
Ile424-Ala433	(1121)	[REDACTED]	
Trp427-Gly431	(1121)	[REDACTED]	
Gln422-Tyr435B	(1121)	[REDACTED]	
Arg426-Gly431	(1121)	[REDACTED]	
Ile423-Met434	(1121)	[REDACTED]	
Gln422-Tyr435	(1121)	[REDACTED]	
Arg426-Lys432	(1121)	[REDACTED]	
Arg426-Gly431B	(1121)	[REDACTED]	
Asn425-Lys432	(1121)	[REDACTED]	
Consensus	(1121)	GCGGCGAGTTCTTACTGCAACAGCACCCAGCTGTTCAA	
	1161		1200
Ile424-Ala433	(1161)	[REDACTED]	
Trp427-Gly431	(1161)	[REDACTED]	
Gln422-Tyr435B	(1161)	[REDACTED]	
Arg426-Gly431	(1161)	[REDACTED]	
Ile423-Met434	(1161)	[REDACTED]	
Gln422-Tyr435	(1161)	[REDACTED]	
Arg426-Lys432	(1161)	[REDACTED]	
Arg426-Gly431B	(1161)	[REDACTED]	
Asn425-Lys432	(1161)	[REDACTED]	
Consensus	(1161)	CAGCACCTGGAAACAACACCATCGGCCCCAACAACACCAAC	
	1201		1240
Ile424-Ala433	(1201)	[REDACTED]	
Trp427-Gly431	(1201)	[REDACTED]	
Gln422-Tyr435B	(1201)	[REDACTED]	
Arg426-Gly431	(1201)	[REDACTED]	
Ile423-Met434	(1201)	[REDACTED]	
Gln422-Tyr435	(1201)	[REDACTED]	
Arg426-Lys432	(1201)	[REDACTED]	
Arg426-Gly431B	(1201)	[REDACTED]	
Asn425-Lys432	(1201)	[REDACTED]	
Consensus	(1201)	GGCACCATCACCCCTGCCCTGCCGCATCAAGCAGATCATCA	
	1241		1280

FIG. 4F

Ile424-Ala433	(1240)	-----S-----G-----C-----G-----	-----S-----G-----C-----G-----
Trp427-Gly431	(1241)	ACCGCTGG-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435B	(1234)	-----C-----C-----C-----C-----	-----C-----C-----C-----C-----
Arg426-Gly431	(1241)	ACCGCGGC-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Ile423-Met434	(1237)	-----G-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435	(1234)	-----G-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Lys432	(1241)	ACCGGGC-----G-----A-----A-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431B	(1241)	ACCGCGGC-----G-----A-----A-----	-----G-----C-----G-----C-----G-----
Asn425-Lys432	(1241)	AC-----G-----C-----C-----A-----	-----G-----C-----G-----C-----G-----
Consensus	(1241)	AC	GGCGGCAAGGCCATGTACGCCCCCCCACCG
		1281	1320
Ile424-Ala433	(1269)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Trp427-Gly431	(1281)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435B	(1257)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431	(1281)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Ile423-Met434	(1263)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435	(1257)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Lys432	(1281)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431B	(1281)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Asn425-Lys432	(1275)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Consensus	(1281)	CGGCCAGATCCGCTGCAGCAGAACATCACCGGCTGCTG	
		1321	1360
Ile424-Ala433	(1309)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Trp427-Gly431	(1321)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435B	(1297)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431	(1321)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Ile423-Met434	(1303)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435	(1297)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Lys432	(1321)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431B	(1321)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Asn425-Lys432	(1315)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Consensus	(1321)	CTGACCCGCGACGGCGGCAAGGAGATCAGAACACCAACCG	
		1361	1400
Ile424-Ala433	(1349)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Trp427-Gly431	(1361)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435B	(1337)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431	(1361)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Ile423-Met434	(1343)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435	(1337)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Lys432	(1361)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431B	(1361)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Asn425-Lys432	(1355)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Consensus	(1361)	AGATCTTCCGCCCCGGCGCGGCGACATGCGCGACAACTG	
		1401	1440
Ile424-Ala433	(1389)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Trp427-Gly431	(1401)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435B	(1377)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431	(1401)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Ile423-Met434	(1383)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Gln422-Tyr435	(1377)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Lys432	(1401)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Arg426-Gly431B	(1401)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Asn425-Lys432	(1395)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Consensus	(1401)	GCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAG	
		1441	1480
Ile424-Ala433	(1429)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----
Trp427-Gly431	(1441)	-----G-----C-----G-----C-----G-----	-----G-----C-----G-----C-----G-----

FIG. 4G

Gln422-Tyr435B	(1417)	[REDACTED]
Arg426-Gly431	(1441)	[REDACTED]
Ile423-Met434	(1423)	[REDACTED]
Gln422-Tyr435	(1417)	[REDACTED]
Arg426-Lys432	(1441)	[REDACTED]
Arg426-Gly431B	(1441)	[REDACTED]
Asn425-Lys432	(1435)	[REDACTED]
Consensus	(1441)	CCCTGGCGTGGCCCCACCAAGCCAAGGCCGCGCGTGG 1481 1520
Ile424-Ala433	(1469)	[REDACTED]
Trp427-Gly431	(1481)	[REDACTED]
Gln422-Tyr435B	(1457)	[REDACTED]
Arg426-Gly431	(1481)	[REDACTED]
Ile423-Met434	(1463)	[REDACTED]
Gln422-Tyr435	(1457)	[REDACTED]
Arg426-Lys432	(1481)	[REDACTED]
Arg426-Gly431B	(1481)	[REDACTED]
Asn425-Lys432	(1475)	[REDACTED]
Consensus	(1481)	TGCAGCGCGAGAACGGCCGTGACCTGGCGCATGTT 1521 1560
Ile424-Ala433	(1509)	[REDACTED]
Trp427-Gly431	(1521)	[REDACTED]
Gln422-Tyr435B	(1497)	[REDACTED]
Arg426-Gly431	(1521)	[REDACTED]
Ile423-Met434	(1503)	[REDACTED]
Gln422-Tyr435	(1497)	[REDACTED]
Arg426-Lys432	(1521)	[REDACTED]
Arg426-Gly431B	(1521)	[REDACTED]
Asn425-Lys432	(1515)	[REDACTED]
Consensus	(1521)	CCTGGGCTTCTGGGCCCCGGCAGCACCATGGCGCC 1561 1600
Ile424-Ala433	(1549)	[REDACTED]
Trp427-Gly431	(1561)	[REDACTED]
Gln422-Tyr435B	(1537)	[REDACTED]
Arg426-Gly431	(1561)	[REDACTED]
Ile423-Met434	(1543)	[REDACTED]
Gln422-Tyr435	(1537)	[REDACTED]
Arg426-Lys432	(1561)	[REDACTED]
Arg426-Gly431B	(1561)	[REDACTED]
Asn425-Lys432	(1555)	[REDACTED]
Consensus	(1561)	CGCAGCCTGACCTGACCGTGCAGGCCCGCCAGCTGCTGA 1601 1640
Ile424-Ala433	(1589)	[REDACTED]
Trp427-Gly431	(1601)	[REDACTED]
Gln422-Tyr435B	(1577)	[REDACTED]
Arg426-Gly431	(1601)	[REDACTED]
Ile423-Met434	(1583)	[REDACTED]
Gln422-Tyr435	(1577)	[REDACTED]
Arg426-Lys432	(1601)	[REDACTED]
Arg426-Gly431B	(1601)	[REDACTED]
Asn425-Lys432	(1595)	[REDACTED]
Consensus	(1601)	GCGGCATCGTGCAGCAGCAGAACACCTGCTGCGGCCAT 1641 1680
Ile424-Ala433	(1629)	[REDACTED]
Trp427-Gly431	(1641)	[REDACTED]
Gln422-Tyr435B	(1617)	[REDACTED]
Arg426-Gly431	(1641)	[REDACTED]

FIG. 4H

FIG. 41

Arg426-Lys432	(1841)	TCCTGGATGGAGTGGGAGCGCGAGATCGACAAC
Arg426-Gly431B	(1841)	TCCTGGATGGAGTGGGAGCGCGAGATCGACAAC
Asn425-Lys432	(1835)	TGACCTGGATGGAGTGGGAGCGCGAGATCGACAAC
Consensus	(1841)	TGACCTGGATGGAGTGGGAGCGCGAGATCGACAAC
	1881	1920
Ile424-Ala433	(1869)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Trp427-Gly431	(1881)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Gln422-Tyr435B	(1857)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Arg426-Gly431	(1881)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Ile423-Met434	(1863)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Gln422-Tyr435	(1857)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Arg426-Lys432	(1881)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Arg426-Gly431B	(1881)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Asn425-Lys432	(1875)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
Consensus	(1881)	GGATGATCTACACCCCTGATCGAGGAGGCCAGAACCAG
	1921	1960
Ile424-Ala433	(1909)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Trp427-Gly431	(1921)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Gln422-Tyr435B	(1897)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Arg426-Gly431	(1921)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Ile423-Met434	(1903)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Gln422-Tyr435	(1897)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Arg426-Lys432	(1921)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Arg426-Gly431B	(1921)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Asn425-Lys432	(1915)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
Consensus	(1921)	CAGGAGAAGAACGGAGCAGGAGCTGCTGGAGCTGGACAAGT
	1961	2000
Ile424-Ala433	(1949)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Trp427-Gly431	(1961)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Gln422-Tyr435B	(1937)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Arg426-Gly431	(1961)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Ile423-Met434	(1943)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Gln422-Tyr435	(1937)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Arg426-Lys432	(1961)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Arg426-Gly431B	(1961)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Asn425-Lys432	(1955)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
Consensus	(1961)	GGGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCT
	2001	2040
Ile424-Ala433	(1989)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Trp427-Gly431	(2001)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Gln422-Tyr435B	(1977)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Arg426-Gly431	(2001)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Ile423-Met434	(1983)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Gln422-Tyr435	(1977)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Arg426-Lys432	(2001)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Arg426-Gly431B	(2001)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Asn425-Lys432	(1995)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Consensus	(2001)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
	2041	2080
Ile424-Ala433	(2029)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Trp427-Gly431	(2041)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Gln422-Tyr435B	(2017)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Arg426-Gly431	(2041)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Ile423-Met434	(2023)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Gln422-Tyr435	(2017)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Arg426-Lys432	(2041)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG
Arg426-Gly431B	(2041)	GTGGTACATCAAGATCTTCATGATCGTGGCGGCCTG

FIG. 4J

Asn425-Lys432	(2035)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Consensus	(2041)	GTGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
	2081	2120
Ile424-Ala433	(2069)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Trp427-Gly431	(2081)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435B	(2057)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431	(2081)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Ile423-Met434	(2063)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435	(2057)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Lys432	(2081)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431B	(2081)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Asn425-Lys432	(2075)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Consensus	(2081)	ACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGAC
	2121	2160
Ile424-Ala433	(2109)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Trp427-Gly431	(2121)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435B	(2097)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431	(2121)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Ile423-Met434	(2103)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435	(2097)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Lys432	(2121)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431B	(2121)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Asn425-Lys432	(2115)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Consensus	(2121)	CCGCTTCCCCGCCCGGGCGGGCCCCGACCGCCCCGAGGGC
	2161	2200
Ile424-Ala433	(2149)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Trp427-Gly431	(2161)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435B	(2137)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431	(2161)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Ile423-Met434	(2143)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435	(2137)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Lys432	(2161)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431B	(2161)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Asn425-Lys432	(2155)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Consensus	(2161)	ATCGAGGAGGGAGGGCGGGAGCGCGACCGCGACCGCAGCA
	2201	2240
Ile424-Ala433	(2189)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Trp427-Gly431	(2201)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435B	(2177)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431	(2201)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Ile423-Met434	(2183)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435	(2177)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Lys432	(2201)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431B	(2201)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Asn425-Lys432	(2195)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Consensus	(2201)	GCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA
	2241	2280
Ile424-Ala433	(2229)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Trp427-Gly431	(2241)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435B	(2217)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431	(2241)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Ile423-Met434	(2223)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Gln422-Tyr435	(2217)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Lys432	(2241)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Arg426-Gly431B	(2241)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Asn425-Lys432	(2235)	ATGGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGA
Consensus	(2241)	CCTGGCAGCCTGTGCCTGTTAGCTACCAACCCTGC

**FIG. 4K**

		2281	2320
Ile424-Ala433	(2269)	[REDACTED]	
Trp427-Gly431	(2281)	[REDACTED]	
Gln422-Tyr435B	(2257)	[REDACTED]	
Arg426-Gly431	(2281)	[REDACTED]	
Ile423-Met434	(2263)	[REDACTED]	
Gln422-Tyr435	(2257)	[REDACTED]	
Arg426-Lys432	(2281)	[REDACTED]	
Arg426-Gly431B	(2281)	[REDACTED]	
Asn425-Lys432	(2275)	[REDACTED]	
Consensus	(2281)	GACCTGATCCTGATGCCGCCGCATCGTGGAGCTGCTGG	
		2321	2360
Ile424-Ala433	(2309)	[REDACTED]	
Trp427-Gly431	(2321)	[REDACTED]	
Gln422-Tyr435B	(2297)	[REDACTED]	
Arg426-Gly431	(2321)	[REDACTED]	
Ile423-Met434	(2303)	[REDACTED]	
Gln422-Tyr435	(2297)	[REDACTED]	
Arg426-Lys432	(2321)	[REDACTED]	
Arg426-Gly431B	(2321)	[REDACTED]	
Asn425-Lys432	(2315)	[REDACTED]	
Consensus	(2321)	GCCGCCGCCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCT	
		2361	2400
Ile424-Ala433	(2349)	[REDACTED]	
Trp427-Gly431	(2361)	[REDACTED]	
Gln422-Tyr435B	(2337)	[REDACTED]	
Arg426-Gly431	(2361)	[REDACTED]	
Ile423-Met434	(2343)	[REDACTED]	
Gln422-Tyr435	(2337)	[REDACTED]	
Arg426-Lys432	(2361)	[REDACTED]	
Arg426-Gly431B	(2361)	[REDACTED]	
Asn425-Lys432	(2355)	[REDACTED]	
Consensus	(2361)	GCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG	
		2401	2440
Ile424-Ala433	(2389)	[REDACTED]	
Trp427-Gly431	(2401)	[REDACTED]	
Gln422-Tyr435B	(2377)	[REDACTED]	
Arg426-Gly431	(2401)	[REDACTED]	
Ile423-Met434	(2383)	[REDACTED]	
Gln422-Tyr435	(2377)	[REDACTED]	
Arg426-Lys432	(2401)	[REDACTED]	
Arg426-Gly431B	(2401)	[REDACTED]	
Asn425-Lys432	(2395)	[REDACTED]	
Consensus	(2401)	AGCCTGTCGACGCCATGCCATGCCGTGGCGAGGGCA	
		2441	2480
Ile424-Ala433	(2429)	[REDACTED]	
Trp427-Gly431	(2441)	[REDACTED]	
Gln422-Tyr435B	(2417)	[REDACTED]	
Arg426-Gly431	(2441)	[REDACTED]	
Ile423-Met434	(2423)	[REDACTED]	
Gln422-Tyr435	(2417)	[REDACTED]	
Arg426-Lys432	(2441)	[REDACTED]	
Arg426-Gly431B	(2441)	[REDACTED]	
Asn425-Lys432	(2435)	[REDACTED]	
Consensus	(2441)	CCGACCGCATCATCGAGGTGGCCCAGCGCATGGCCCGCG	
		2481	2520
Ile424-Ala433	(2469)	[REDACTED]	

FIG. 4L

Trp427-Gly431	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Gln422-Tyr435B	(2457)	CGTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Arg426-Gly431	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Ile423-Met434	(2463)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Gln422-Tyr435	(2457)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Arg426-Lys432	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Arg426-Gly431B	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Asn425-Lys432	(2475)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Consensus	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
		2521 2541
Ile424-Ala433	(2509)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Trp427-Gly431	(2521)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Gln422-Tyr435B	(2497)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Arg426-Gly431	(2521)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Ile423-Met434	(2503)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Gln422-Tyr435	(2497)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Arg426-Lys432	(2521)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Arg426-Gly431B	(2521)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Asn425-Lys432	(2515)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
Consensus	(2521)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG

FIG. 4M

28 / 65

Leu122-Ser199-Tryp427-Gly431	(1) GAATTGCCACCATGGATGCAATGAAGAGA	30
Val1127-Asn195-Arg426-Gly431	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Val1120-Thr202-Ile424-Ala433	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Leu122-Ser199-Arg426-Lys432	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Leu122-Ser199-Arg426-Gly431	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Lys121-Val200-Asn425-Lys432	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Val1120-Ile201-Ile424-Ala433	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Val1120-Ile201B-Ile424-Ala433	(1) GAATTGCCACCATGGATGCAATGAAGAGA	
Consensus	31	60
Leu122-Ser199-Tryp427-Gly431	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Val1127-Asn195-Arg426-Gly431	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Val1120-Thr202-Ile424-Ala433	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Leu122-Ser199-Arg426-Lys432	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Leu122-Ser199-Arg426-Gly431	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Lys121-Val200-Asn425-Lys432	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Val1120-Ile201-Ile424-Ala433	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Val1120-Ile201B-Ile424-Ala433	(31) GGGCTCTGCTGTGCTGCTGCTGTGCTGGA	
Consensus	61	90
Leu122-Ser199-Tryp427-Gly431	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Val1127-Asn195-Arg426-Gly431	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Val1120-Thr202-Ile424-Ala433	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Leu122-Ser199-Arg426-Lys432	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Leu122-Ser199-Arg426-Gly431	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Lys121-Val200-Asn425-Lys432	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Val1120-Ile201-Ile424-Ala433	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Val1120-Ile201B-Ile424-Ala433	(61) GCAGTCTTCGTTTCGCCAGGCCGTGGAG	
Consensus	91	120
Leu122-Ser199-Tryp427-Gly431	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Val1127-Asn195-Arg426-Gly431	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Val1120-Thr202-Ile424-Ala433	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Leu122-Ser199-Arg426-Lys432	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Leu122-Ser199-Arg426-Gly431	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Lys121-Val200-Asn425-Lys432	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Val1120-Ile201-Ile424-Ala433	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Val1120-Ile201B-Ile424-Ala433	(91) AAGCTGTGGGTGACCGTGTACTACGGCGTG	
Consensus	121	150
Leu122-Ser199-Tryp427-Gly431	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Val1127-Asn195-Arg426-Gly431	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Val1120-Thr202-Ile424-Ala433	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Leu122-Ser199-Arg426-Lys432	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Leu122-Ser199-Arg426-Gly431	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Lys121-Val200-Asn425-Lys432	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Val1120-Ile201-Ile424-Ala433	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Val1120-Ile201B-Ile424-Ala433	(121) CCCGTGTGGAAGGAGGCCACCAACCGCTG	
Consensus	151	180
Leu122-Ser199-Tryp427-Gly431	(151) TTCTGCGCCAGCGACGCCAAGGCCCTACGAC	
Val1127-Asn195-Arg426-Gly431	(151) TTCTGCGCCAGCGACGCCAAGGCCCTACGAC	
Val1120-Thr202-Ile424-Ala433	(151) TTCTGCGCCAGCGACGCCAAGGCCCTACGAC	
Leu122-Ser199-Arg426-Lys432	(151) TTCTGCGCCAGCGACGCCAAGGCCCTACGAC	
Leu122-Ser199-Arg426-Gly431	(151) TTCTGCGCCAGCGACGCCAAGGCCCTACGAC	
Lys121-Val200-Asn425-Lys432	(151) TTCTGCGCCAGCGACGCCAAGGCCCTACGAC	

Val120-Ile201-Ile424-Ala433	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Val120-Ile201B-Ile424-Ala433	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Consensus	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
	181	210
Leu122-Ser199-Tryp427-Gly431	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Val127-Asn195-Arg426-Gly431	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Val120-Thr202-Ile424-Ala433	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Leu122-Ser199-Arg426-Lys432	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Leu122-Ser199-Arg426-Gly431	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Lys121-Val200-Asn425-Lys432	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Val120-Ile201-Ile424-Ala433	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Val120-Ile201B-Ile424-Ala433	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
Consensus	(181)	ACCGAGGTGCACAACGTGTGGCCACCCAC
	211	240
Leu122-Ser199-Tryp427-Gly431	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Val127-Asn195-Arg426-Gly431	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Val120-Thr202-Ile424-Ala433	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Leu122-Ser199-Arg426-Lys432	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Leu122-Ser199-Arg426-Gly431	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Lys121-Val200-Asn425-Lys432	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Val120-Ile201-Ile424-Ala433	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Val120-Ile201B-Ile424-Ala433	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
Consensus	(211)	GCCTGCGTCCCCACCGACCCCAACCCCCAG
	241	270
Leu122-Ser199-Tryp427-Gly431	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Val127-Asn195-Arg426-Gly431	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Val120-Thr202-Ile424-Ala433	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Leu122-Ser199-Arg426-Lys432	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Leu122-Ser199-Arg426-Gly431	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Lys121-Val200-Asn425-Lys432	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Val120-Ile201-Ile424-Ala433	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Val120-Ile201B-Ile424-Ala433	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
Consensus	(241)	GAGATCGTGTGGAGAACGTGACCGAGAAC
	271	300
Leu122-Ser199-Tryp427-Gly431	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Val127-Asn195-Arg426-Gly431	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Val120-Thr202-Ile424-Ala433	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Leu122-Ser199-Arg426-Lys432	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Leu122-Ser199-Arg426-Gly431	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Lys121-Val200-Asn425-Lys432	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Val120-Ile201-Ile424-Ala433	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Val120-Ile201B-Ile424-Ala433	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
Consensus	(271)	TTCAACATGTGGAGAACATCATGGTGGAG
	301	330
Leu122-Ser199-Tryp427-Gly431	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Val127-Asn195-Arg426-Gly431	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Val120-Thr202-Ile424-Ala433	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Leu122-Ser199-Arg426-Lys432	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Leu122-Ser199-Arg426-Gly431	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Lys121-Val200-Asn425-Lys432	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Val120-Ile201-Ile424-Ala433	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Val120-Ile201B-Ile424-Ala433	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
Consensus	(301)	CAGATGCAGGAGGACATCATCAGGCTGTGG
	331	360
Leu122-Ser199-Tryp427-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Val127-Asn195-Arg426-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Val120-Thr202-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGT-----

	30	/	65	
Leu122-Ser199-Arg426-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG		
Leu122-Ser199-Arg426-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG		
Lys121-Val200-Asn425-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG		
Val120-Ile201-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG		
Val120-Ile201B-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG		
Consensus		361		390
Leu122-Ser199-Tryp427-Gly431	(361)	-----GG-----		
Val127-Asn195-Arg426-Gly431	(361)	ACCCCCCTGTGCGTGGGGCAGGGAACTGC		
Val120-Thr202-Ile424-Ala433	(355)	-----GG-----		
Leu122-Ser199-Arg426-Lys432	(361)	-----GG-----		
Leu122-Ser199-Arg426-Gly431	(361)	-----GG-----		
Lys121-Val200-Asn425-Lys432	(357)	-----GG-----		
Val120-Ile201-Ile424-Ala433	(355)	-----		
Val120-Ile201B-Ile424-Ala433	(355)	-----		
Consensus	(361)	391		420
Leu122-Ser199-Tryp427-Gly431	(363)	--CAACAGCGTGTATCACCCAGGCCTGCC		
Val127-Asn195-Arg426-Gly431	(391)	AAACACCGGGTGTATCACCCAGGCCTGCC		
Val120-Thr202-Ile424-Ala433	(357)	-----GGGGC---CACCCAGGCCTGCC		
Leu122-Ser199-Arg426-Lys432	(363)	--CAACAGCGTGTATCACCCAGGCCTGCC		
Leu122-Ser199-Arg426-Gly431	(363)	--CAACAGCGTGTATCACCCAGGCCTGCC		
Lys121-Val200-Asn425-Lys432	(359)	-----CECCCGGTGTATCACCCAGGCCTGCC		
Val120-Ile201-Ile424-Ala433	(355)	-----GLCGGCATCACCCAGGCCTGCC		
Val120-Ile201B-Ile424-Ala433	(355)	-----CCCGGCATCACCCAGGCCTGCC		
Consensus	(391)	CA CAGCGTGTATCACCCAGGCCTGCC		
Leu122-Ser199-Tryp427-Gly431	(391)	421		450
Val127-Asn195-Arg426-Gly431	(421)	AAGCTGAGCTTGGAGCCATCCCCATCCAC		
Val120-Thr202-Ile424-Ala433	(379)	AAGGTGAGCTTGGAGCCATCCCCATCCAC		
Leu122-Ser199-Arg426-Lys432	(391)	AAGCTGAGCTTGGAGCCATCCCCATCCAC		
Leu122-Ser199-Arg426-Gly431	(391)	AAGCTGAGCTTGGAGCCATCCCCATCCAC		
Lys121-Val200-Asn425-Lys432	(385)	AAGGTGAGCTTGGAGCCATCCCCATCCAC		
Val120-Ile201-Ile424-Ala433	(379)	AAGCTGAGCTTGGAGCCATCCCCATCCAC		
Val120-Ile201B-Ile424-Ala433	(379)	AAGCTGAGCTTGGAGCCATCCCCATCCAC		
Consensus	(421)	AAGGTGAGCTTGGAGCCATCCCCATCCAC		
Leu122-Ser199-Tryp427-Gly431	(421)	451		480
Val127-Asn195-Arg426-Gly431	(451)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Val120-Thr202-Ile424-Ala433	(409)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Leu122-Ser199-Arg426-Lys432	(421)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Leu122-Ser199-Arg426-Gly431	(421)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Lys121-Val200-Asn425-Lys432	(415)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Val120-Ile201-Ile424-Ala433	(409)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Val120-Ile201B-Ile424-Ala433	(409)	TACTGGCGGGGGGGCTTGGGCATGGTG		
Consensus	(451)	TACTGCGCCCCCGCCGGCTTCGCCATGGTG		
Leu122-Ser199-Tryp427-Gly431	(451)	481		510
Val127-Asn195-Arg426-Gly431	(481)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Val120-Thr202-Ile424-Ala433	(439)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Leu122-Ser199-Arg426-Lys432	(451)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Leu122-Ser199-Arg426-Gly431	(451)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Lys121-Val200-Asn425-Lys432	(445)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Val120-Ile201-Ile424-Ala433	(439)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Val120-Ile201B-Ile424-Ala433	(439)	AAGTCCAACGACAAAGAAGTTCAACGGGAGC		
Consensus	(481)	511		540

	31	/	65
Leu122-Ser199-Tryp427-Gly431	(481)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Val127-Asn195-Arg426-Gly431	(511)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Val120-Thr202-Ile424-Ala433	(469)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Leu122-Ser199-Arg426-Lys432	(481)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Leu122-Ser199-Arg426-Gly431	(481)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Lys121-Val200-Asn425-Lys432	(475)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Val120-Ile201-Ile424-Ala433	(469)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Val120-Ile201B-Ile424-Ala433	(469)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	
Consensus	(511)	GGCCCCCTGCACCAACGTGAGCACCCTGCAG	570
Leu122-Ser199-Tryp427-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Val127-Asn195-Arg426-Gly431	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Val120-Thr202-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Leu122-Ser199-Arg426-Lys432	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Leu122-Ser199-Arg426-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Lys121-Val200-Asn425-Lys432	(505)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Val120-Ile201-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Val120-Ile201B-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	
Consensus	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC	600
Leu122-Ser199-Tryp427-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val127-Asn195-Arg426-Gly431	(571)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val120-Thr202-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Leu122-Ser199-Arg426-Lys432	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Leu122-Ser199-Arg426-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Lys121-Val200-Asn425-Lys432	(535)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val120-Ile201-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val120-Ile201B-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Consensus	(571)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC	630
Leu122-Ser199-Tryp427-Gly431	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Val127-Asn195-Arg426-Gly431	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Val120-Thr202-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Leu122-Ser199-Arg426-Lys432	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Leu122-Ser199-Arg426-Gly431	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Lys121-Val200-Asn425-Lys432	(565)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Val120-Ile201-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Val120-Ile201B-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	
Consensus	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC	660
Leu122-Ser199-Tryp427-Gly431	(601)	TTCACCGACAAACGCCAACGACATCATCGTG	
Val127-Asn195-Arg426-Gly431	(631)	TTCACCGACAAACGCCAACGACATCATCGTG	
Val120-Thr202-Ile424-Ala433	(589)	TTCACCGACAAACGCCAACGACATCATCGTG	
Leu122-Ser199-Arg426-Lys432	(601)	TTCACCGACAAACGCCAACGACATCATCGTG	
Leu122-Ser199-Arg426-Gly431	(601)	TTCACCGACAAACGCCAACGACATCATCGTG	
Lys121-Val200-Asn425-Lys432	(595)	TTCACCGACAAACGCCAACGACATCATCGTG	
Val120-Ile201-Ile424-Ala433	(589)	TTCACCGACAAACGCCAACGACATCATCGTG	
Val120-Ile201B-Ile424-Ala433	(589)	TTCACCGACAAACGCCAACGACATCATCGTG	
Consensus	(631)	TTCACCGACAAACGCCAACGACATCATCGTG	690
Leu122-Ser199-Tryp427-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Val127-Asn195-Arg426-Gly431	(661)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Val120-Thr202-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Leu122-Ser199-Arg426-Lys432	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Leu122-Ser199-Arg426-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Lys121-Val200-Asn425-Lys432	(625)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Val120-Ile201-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC	

	32	/	65	
Val120-Ile201B-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC		
Consensus	(661)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC		
		691		720
Leu122-Ser199-Tryp427-Gly431	(661)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Val127-Asn195-Arg426-Gly431	(691)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Val120-Thr202-Ile424-Ala433	(649)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Leu122-Ser199-Arg426-Lys432	(661)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Leu122-Ser199-Arg426-Gly431	(661)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Lys121-Val200-Asn425-Lys432	(655)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Val120-Ile201-Ile424-Ala433	(649)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Val120-Ile201B-Ile424-Ala433	(649)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
Consensus	(691)	ACCCGGCCCCAACAAACAACACCCGCAAGAGC		
		721		750
Leu122-Ser199-Tryp427-Gly431	(691)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Val127-Asn195-Arg426-Gly431	(721)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Val120-Thr202-Ile424-Ala433	(679)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Leu122-Ser199-Arg426-Lys432	(691)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Leu122-Ser199-Arg426-Gly431	(691)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Lys121-Val200-Asn425-Lys432	(685)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Val120-Ile201-Ile424-Ala433	(679)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Val120-Ile201B-Ile424-Ala433	(679)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
Consensus	(721)	ATCACCATCGGCCCCGGCGCGCCTCTAC		
		751		780
Leu122-Ser199-Tryp427-Gly431	(721)	GCCACCGGGGACATCATCGGCGACATCCGC		
Val127-Asn195-Arg426-Gly431	(751)	GCCACCGGGGACATCATCGGCGACATCCGC		
Val120-Thr202-Ile424-Ala433	(709)	GCCACCGGGGACATCATCGGCGACATCCGC		
Leu122-Ser199-Arg426-Lys432	(721)	GCCACCGGGGACATCATCGGCGACATCCGC		
Leu122-Ser199-Arg426-Gly431	(721)	GCCACCGGGGACATCATCGGCGACATCCGC		
Lys121-Val200-Asn425-Lys432	(715)	GCCACCGGGGACATCATCGGCGACATCCGC		
Val120-Ile201-Ile424-Ala433	(709)	GCCACCGGGGACATCATCGGCGACATCCGC		
Val120-Ile201B-Ile424-Ala433	(709)	GCCACCGGGGACATCATCGGCGACATCCGC		
Consensus	(751)	GCCACCGGGGACATCATCGGCGACATCCGC		
		781		810
Leu122-Ser199-Tryp427-Gly431	(751)	CAGCCCCACTGAAACATCAGCGGGGAGAAG		
Val127-Asn195-Arg426-Gly431	(781)	CAGCCCCACTGAAACATCAGCGGGGAGAAG		
Val120-Thr202-Ile424-Ala433	(739)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
Leu122-Ser199-Arg426-Lys432	(751)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
Leu122-Ser199-Arg426-Gly431	(751)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
Lys121-Val200-Asn425-Lys432	(745)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
Val120-Ile201-Ile424-Ala433	(739)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
Val120-Ile201B-Ile424-Ala433	(739)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
Consensus	(781)	CAGGGCCACTGAAACATCAGCGGGGAGAAG		
		811		840
Leu122-Ser199-Tryp427-Gly431	(781)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Val127-Asn195-Arg426-Gly431	(811)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Val120-Thr202-Ile424-Ala433	(769)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Leu122-Ser199-Arg426-Lys432	(781)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Leu122-Ser199-Arg426-Gly431	(781)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Lys121-Val200-Asn425-Lys432	(775)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Val120-Ile201-Ile424-Ala433	(769)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Val120-Ile201B-Ile424-Ala433	(769)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
Consensus	(811)	TGGAACAAACACCCCTGAAGCAGATCGTGACC		
		841		870
Leu122-Ser199-Tryp427-Gly431	(811)	AAGCTGCAGGGCCCAGTTGGCAACAAAGACC		
Val127-Asn195-Arg426-Gly431	(841)	AAGCTGCAGGGCCCAGTTGGCAACAAAGACC		
Val120-Thr202-Ile424-Ala433	(799)	AAGCTGCAGGGCCCAGTTGGCAACAAAGACC		
Leu122-Ser199-Arg426-Lys432	(811)	AAGCTGCAGGGCCCAGTTGGCAACAAAGACC		

	33	/	65
Leu122-Ser199-Arg426-Gly431	(811)	AAGCTGCAGGCCAGTTCGGCAACAAGACC	
Lys121-Val200-Asn425-Lys432	(805)	AAGCTGCAGGCCAGTTCGGCAACAAGACC	
Val120-Ile201-Ile424-Ala433	(799)	AAGCTGCAGGCCAGTTCGGCAACAAGACC	
Val120-Ile201B-Ile424-Ala433	(799)	AAGCTGCAGGCCAGTTCGGCAACAAGACC	
Consensus	(841)	AAGCTGCAGGCCAGTTCGGCAACAAGACC	
		871	900
Leu122-Ser199-Tryp427-Gly431	(841)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Val127-Asn195-Arg426-Gly431	(871)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Val120-Thr202-Ile424-Ala433	(829)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Leu122-Ser199-Arg426-Lys432	(841)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Leu122-Ser199-Arg426-Gly431	(841)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Lys121-Val200-Asn425-Lys432	(835)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Val120-Ile201-Ile424-Ala433	(829)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Val120-Ile201B-Ile424-Ala433	(829)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
Consensus	(871)	ATCGTGTTCAGCAGAGCAGCGGCCGAC	
		901	930
Leu122-Ser199-Tryp427-Gly431	(871)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Val127-Asn195-Arg426-Gly431	(901)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Val120-Thr202-Ile424-Ala433	(859)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Leu122-Ser199-Arg426-Lys432	(871)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Leu122-Ser199-Arg426-Gly431	(871)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Lys121-Val200-Asn425-Lys432	(865)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Val120-Ile201-Ile424-Ala433	(859)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Val120-Ile201B-Ile424-Ala433	(859)	CCCGAGATCGTGTGACAGCTTCAACTGC	
Consensus	(901)	CCCGAGATCGTGTGACAGCTTCAACTGC	
		931	960
Leu122-Ser199-Tryp427-Gly431	(901)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Val127-Asn195-Arg426-Gly431	(931)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Val120-Thr202-Ile424-Ala433	(889)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Leu122-Ser199-Arg426-Lys432	(901)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Leu122-Ser199-Arg426-Gly431	(901)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Lys121-Val200-Asn425-Lys432	(895)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Val120-Ile201-Ile424-Ala433	(889)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Val120-Ile201B-Ile424-Ala433	(889)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
Consensus	(931)	GGGGGGGAGTCTTCTACTGCAACAGCACC	
		961	990
Leu122-Ser199-Tryp427-Gly431	(931)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Val127-Asn195-Arg426-Gly431	(961)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Val120-Thr202-Ile424-Ala433	(919)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Leu122-Ser199-Arg426-Lys432	(931)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Leu122-Ser199-Arg426-Gly431	(931)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Lys121-Val200-Asn425-Lys432	(925)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Val120-Ile201-Ile424-Ala433	(919)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Val120-Ile201B-Ile424-Ala433	(919)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
Consensus	(961)	CAGCTGTTAACAGGACCTGGAAACAAGCC	
		991	1020
Leu122-Ser199-Tryp427-Gly431	(961)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Val127-Asn195-Arg426-Gly431	(991)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Val120-Thr202-Ile424-Ala433	(949)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Leu122-Ser199-Arg426-Lys432	(961)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Leu122-Ser199-Arg426-Gly431	(961)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Lys121-Val200-Asn425-Lys432	(955)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Val120-Ile201-Ile424-Ala433	(949)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Val120-Ile201B-Ile424-Ala433	(949)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
Consensus	(991)	ATCGGGCCCCAACAAACACCAACGGCACCATC	
		1021	1050
Leu122-Ser199-Tryp427-Gly431	(991)	ACCTCTGCCCTGCCGCATCAAGCAGATCATC	

Val127-Asn195-Arg426-Gly431	(1021)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Val120-Thr202-Ile424-Ala433	(979)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Leu122-Ser199-Arg426-Lys432	(991)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Leu122-Ser199-Arg426-Gly431	(991)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Lys121-Val200-Asn425-Lys432	(985)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Val120-Ile201-Ile424-Ala433	(979)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Val120-Ile201B-Ile424-Ala433	(979)	ACCTGCGCTGCCGCATCAAGCAGATCATC
Consensus	(1021)	ACCTGCGCTGCCGCATCAAGCAGATCATC
		1051
Leu122-Ser199 Tryp427-Gly431	(1021)	AACCGCTGGGCGGGCAAGGCCATGTACGCC
Val127-Asn195-Arg426-Gly431	(1051)	AACCGCGGCGGGCAAGGCCATGTACGCC
Val120-Thr202-Ile424-Ala433	(1009)	-----GGCGGG-----GCCATGTACGCC
Leu122-Ser199-Arg426-Lys432	(1021)	AACCGGGCGCAAGGCCATGTACGCC
Leu122-Ser199-Arg426-Gly431	(1021)	AACCGCGGCAAGGCCATGTACGCC
Lys121-Val200-Asn425-Lys432	(1015)	AAC-----GCCCAAGGCCATGTACGCC
Val120-Ile201-Ile424-Ala433	(1009)	-----GGCGGC-----GCCATGTACGCC
Val120-Ile201B-Ile424-Ala433	(1009)	-----GGCGGC-----GCCATGTACGCC
Consensus	(1051)	AACCGC G GGCGGCAAGGCCATGTACGCC
		1081
Leu122-Ser199 Tryp427-Gly431	(1051)	CCCCCGATCCGGGGCCAGATCCGCTGCAGC
Val127-Asn195-Arg426-Gly431	(1081)	CCCCCGATCCGGGGCCAGATCCGCTGCAGC
Val120-Thr202-Ile424-Ala433	(1027)	CCCCCGATCCGGGGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Lys432	(1051)	CCCCCGATCCGGGGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Gly431	(1051)	CCCCCGATCCGGGGCCAGATCCGCTGCAGC
Lys121-Val200-Asn425-Lys432	(1039)	CCCCCCATCCGGGGCCAGATCCGCTGCAGC
Val120-Ile201-Ile424-Ala433	(1027)	CCCCCCATCCGGGGCCAGATCCGCTGCAGC
Val120-Ile201B-Ile424-Ala433	(1027)	CCCCCCATCCGGGGCCAGATCCGCTGCAGC
Consensus	(1081)	CCCCCCATCCGGGGCCAGATCCGCTGCAGC
		1111
Leu122-Ser199 Tryp427-Gly431	(1081)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Val127-Asn195-Arg426-Gly431	(1111)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Val120-Thr202-Ile424-Ala433	(1057)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Leu122-Ser199-Arg426-Lys432	(1081)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Leu122-Ser199-Arg426-Gly431	(1081)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Lys121-Val200-Asn425-Lys432	(1069)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Val120-Ile201-Ile424-Ala433	(1057)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Val120-Ile201B-Ile424-Ala433	(1057)	AGCAACATCACCGGCTGGCTGCTGACCCGC
Consensus	(1111)	AGCAACATCACCGGCTGGCTGCTGACCCGC
		1141
Leu122-Ser199 Tryp427-Gly431	(1111)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Val127-Asn195-Arg426-Gly431	(1141)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Val120-Thr202-Ile424-Ala433	(1087)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Leu122-Ser199-Arg426-Lys432	(1111)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Leu122-Ser199-Arg426-Gly431	(1111)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Lys121-Val200-Asn425-Lys432	(1099)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Val120-Ile201-Ile424-Ala433	(1087)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Val120-Ile201B-Ile424-Ala433	(1087)	GACGGGGGGCAAGGAGATCAGCAACACCAC
Consensus	(1141)	GACGGGGGGCAAGGAGATCAGCAACACCAC
		1171
Leu122-Ser199 Tryp427-Gly431	(1141)	GAGATCTTCCGGCCCGGGCGGGGAGATG
Val127-Asn195-Arg426-Gly431	(1171)	GAGATCTTCCGGCCCGGGCGGGGAGATG
Val120-Thr202-Ile424-Ala433	(1117)	GAGATCTTCCGGCCCGGGCGGGGAGATG
Leu122-Ser199-Arg426-Lys432	(1141)	GAGATCTTCCGGCCCGGGCGGGGAGATG
Leu122-Ser199-Arg426-Gly431	(1141)	GAGATCTTCCGGCCCGGGCGGGGAGATG
Lys121-Val200-Asn425-Lys432	(1129)	GACATCTTCCGGCCCGGGCGGGGAGATG
Val120-Ile201-Ile424-Ala433	(1117)	GAGATCTTCCGGCCCGGGCGGGGAGATG
Val120-Ile201B-Ile424-Ala433	(1117)	GAGATCTTCCGGCCCGGGCGGGGAGATG

Consensus	(1171)	GAGATCTTCCGCCCGGGCGGCACATG
	1201	1230
Leu122-Ser199 Tryp427-Gly431	(1171)	CGCGCATCTCTCCGAGCTGACATG
Val127-Asn195-Arg426-Gly431	(1201)	CGCGCACATGCGATGAGGAGCTACAG
Val120-Thr202-Ile424-Ala433	(1147)	CGCGCACATGCGATGAGGAGCTACAG
Leu122-Ser199-Arg426-Lys432	(1171)	CGCGCACATGCGATGAGGAGCTACAG
Leu122-Ser199-Arg426-Gly431	(1171)	CGCGCACATGCGATGAGGAGCTACAG
Lys121-Val200-Asn425-Lys432	(1159)	CGCGCACATGCGATGAGGAGCTACAG
Val120-Ile201-Ile424-Ala433	(1147)	CGCGCACATGCGATGAGGAGCTACAG
Val120-Ile201B-Ile424-Ala433	(1147)	CGCGCACATGCGATGAGGAGCTACAG
Consensus	(1201)	CGCGACAACTGGCGCAGCGAGCTGACAG
	1231	1260
Leu122-Ser199 Tryp427-Gly431	(1201)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Val127-Asn195-Arg426-Gly431	(1231)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Val120-Thr202-Ile424-Ala433	(1177)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Leu122-Ser199-Arg426-Lys432	(1201)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Leu122-Ser199-Arg426-Gly431	(1201)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Lys121-Val200-Asn425-Lys432	(1189)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Val120-Ile201-Ile424-Ala433	(1177)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Val120-Ile201B-Ile424-Ala433	(1177)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
Consensus	(1231)	TACAAAGGTGGTGAAGATCGAGCCCTGGGC
	1261	1290
Leu122-Ser199 Tryp427-Gly431	(1231)	GTGGCCCCCACCAAGGCCAAGGGCGCGCGTG
Val127-Asn195-Arg426-Gly431	(1261)	GTGGCCCCCACCAAGGCCAAGGGCGCGCGTG
Val120-Thr202-Ile424-Ala433	(1207)	GTGGGGCCCAAGGCCAAGGGCGCGCGTG
Leu122-Ser199-Arg426-Lys432	(1231)	GTGGGGCCCAAGGCCAAGGGCGCGCGTG
Leu122-Ser199-Arg426-Gly431	(1231)	GTGGGGCCCAAGGCCAAGGGCGCGCGTG
Lys121-Val200-Asn425-Lys432	(1219)	GTGGGGCCCAAGGCCAAGGGCGCGCGTG
Val120-Ile201-Ile424-Ala433	(1207)	GTGGGGCCCAAGGCCAAGGGCGCGCGTG
Val120-Ile201B-Ile424-Ala433	(1207)	GTGGGGCCCAAGGCCAAGGGCGCGCGTG
Consensus	(1261)	GTGGCCCCCACCAAGGCCAAGGGCGCGCGTG
	1291	1320
Leu122-Ser199 Tryp427-Gly431	(1261)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Val127-Asn195-Arg426-Gly431	(1291)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Val120-Thr202-Ile424-Ala433	(1237)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Leu122-Ser199-Arg426-Lys432	(1261)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Leu122-Ser199-Arg426-Gly431	(1261)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Lys121-Val200-Asn425-Lys432	(1249)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Val120-Ile201-Ile424-Ala433	(1237)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Val120-Ile201B-Ile424-Ala433	(1237)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
Consensus	(1291)	GTGCAAGGGCTAGAACGGCGCGCGTGACCCCTG
	1321	1350
Leu122-Ser199 Tryp427-Gly431	(1291)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Val127-Asn195-Arg426-Gly431	(1321)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Val120-Thr202-Ile424-Ala433	(1267)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Leu122-Ser199-Arg426-Lys432	(1291)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Leu122-Ser199-Arg426-Gly431	(1291)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Lys121-Val200-Asn425-Lys432	(1279)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Val120-Ile201-Ile424-Ala433	(1267)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Val120-Ile201B-Ile424-Ala433	(1267)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Consensus	(1321)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
	1351	1380
Leu122-Ser199 Tryp427-Gly431	(1321)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Val127-Asn195-Arg426-Gly431	(1351)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Val120-Thr202-Ile424-Ala433	(1297)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Leu122-Ser199-Arg426-Lys432	(1321)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG
Leu122-Ser199-Arg426-Gly431	(1321)	GGCGCCATGTTGGCTTGGCGTGGCGGGCG

Lys121-Val1200-Asn425-Lys432	(1309)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Val120-Ile201-Ile424-Ala433	(1297)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Val120-Ile201B-Ile424-Ala433	(1297)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Consensus	(1351)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
	1381	1410
Leu122-Ser199 Tryp427-Gly431	(1351)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Val127-Asn195-Arg426-Gly431	(1381)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Val120-Thr202-Ile424-Ala433	(1327)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(1351)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(1351)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Lys121-Val1200-Asn425-Lys432	(1339)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Val120-Ile201-Ile424-Ala433	(1327)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(1327)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
Consensus	(1381)	ACCCCTGACCGTGAGGCCCGCCAGCTGCTG
	1411	1440
Leu122-Ser199 Tryp427-Gly431	(1381)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Val127-Asn195-Arg426-Gly431	(1411)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Val120-Thr202-Ile424-Ala433	(1357)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Leu122-Ser199-Arg426-Lys432	(1381)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Leu122-Ser199-Arg426-Gly431	(1381)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Lys121-Val1200-Asn425-Lys432	(1369)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Val120-Ile201-Ile424-Ala433	(1357)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Val120-Ile201B-Ile424-Ala433	(1357)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
Consensus	(1411)	AGCGGCATCGTGAGCAGCAGAACAAACCTG
	1441	1470
Leu122-Ser199 Tryp427-Gly431	(1411)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Val127-Asn195-Arg426-Gly431	(1441)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Val120-Thr202-Ile424-Ala433	(1387)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Leu122-Ser199-Arg426-Lys432	(1411)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Leu122-Ser199-Arg426-Gly431	(1411)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Lys121-Val1200-Asn425-Lys432	(1399)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Val120-Ile201-Ile424-Ala433	(1387)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Val120-Ile201B-Ile424-Ala433	(1387)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
Consensus	(1441)	CTGGCGGCCATCGAGGCCAGCAGCACCTG
	1471	1500
Leu122-Ser199 Tryp427-Gly431	(1441)	CTGGGCTGAGGGTGTTGGGCATCAAGCAG
Val127-Asn195-Arg426-Gly431	(1471)	CTGGGCTGAGGGTGTTGGGCATCAAGCAG
Val120-Thr202-Ile424-Ala433	(1417)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
Leu122-Ser199-Arg426-Lys432	(1441)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
Leu122-Ser199-Arg426-Gly431	(1441)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
Lys121-Val1200-Asn425-Lys432	(1429)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
Val120-Ile201-Ile424-Ala433	(1417)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
Val120-Ile201B-Ile424-Ala433	(1417)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
Consensus	(1471)	CTGCACCTGAGGGTGTTGGGCATCAAGCAG
	1501	1530
Leu122-Ser199 Tryp427-Gly431	(1471)	CTGCAGCTGAGGGTGCTGGGGCATCAAGCAG
Val127-Asn195-Arg426-Gly431	(1501)	CTGCAGCTGAGGGTGCTGGGGCATCAAGCAG
Val120-Thr202-Ile424-Ala433	(1447)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
Leu122-Ser199-Arg426-Lys432	(1471)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
Leu122-Ser199-Arg426-Gly431	(1471)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
Lys121-Val1200-Asn425-Lys432	(1459)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
Val120-Ile201-Ile424-Ala433	(1447)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
Val120-Ile201B-Ile424-Ala433	(1447)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
Consensus	(1501)	CTGCAGGGCCCGGGTGCTGGGGCATCAAGCAG
	1531	1560
Leu122-Ser199 Tryp427-Gly431	(1501)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Val127-Asn195-Arg426-Gly431	(1531)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC

Val120-Thr202-Ile424-Ala433	(1477)	TACCTGAAGGACCCAGGAGCTGCTGGCATC
Leu122-Ser199-Arg426-Lys432	(1501)	TACCTGAGGACCCAGGAGCTGCTGGCATC
Leu122-Ser199-Arg426-Gly431	(1501)	TACCTGAAGGACCCAGGAGCTGCTGGCATC
Lys121-Val200-Asn425-Lys432	(1489)	TACCTGAAGGACCCAGGAGCTGCTGGCATC
Val120-Ile201-Ile424-Ala433	(1477)	TACCTGAGGACCCAGGAGCTGCTGGCATC
Val120-Ile201B-Ile424-Ala433	(1477)	TACCTGAAGGACCCAGGAGCTGCTGGCATC
Consensus	(1531)	TACCTGAAGGACCCAGGAGCTGCTGGCATC
	1561	1590
Leu122-Ser199 Tryp427-Gly431	(1531)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Val127-Asn195-Arg426-Gly431	(1561)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Val120-Thr202-Ile424-Ala433	(1507)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Leu122-Ser199-Arg426-Lys432	(1531)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Leu122-Ser199-Arg426-Gly431	(1531)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Lys121-Val200-Asn425-Lys432	(1519)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Val120-Ile201-Ile424-Ala433	(1507)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Val120-Ile201B-Ile424-Ala433	(1507)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
Consensus	(1561)	TGGGGCTGCAGGGCAAGCTGATCTGCACC
	1591	1620
Leu122-Ser199 Tryp427-Gly431	(1561)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Val127-Asn195-Arg426-Gly431	(1591)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Val120-Thr202-Ile424-Ala433	(1537)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Leu122-Ser199-Arg426-Lys432	(1561)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Leu122-Ser199-Arg426-Gly431	(1561)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Lys121-Val200-Asn425-Lys432	(1549)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Val120-Ile201-Ile424-Ala433	(1537)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Val120-Ile201B-Ile424-Ala433	(1537)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
Consensus	(1591)	ACCCGGCTGCCTGGAAACGCCAGCTGGAGC
	1621	1650
Leu122-Ser199 Tryp427-Gly431	(1591)	AACAAAGGCCCTGGACAGATACTGGAACATC
Val127-Asn195-Arg426-Gly431	(1621)	AACAAAGGCCCTGGACAGATACTGGAACATC
Val120-Thr202-Ile424-Ala433	(1567)	AACAAAGGCCCTGGACAGATACTGGAACATC
Leu122-Ser199-Arg426-Lys432	(1591)	AACAAAGGCCCTGGACAGATACTGGAACATC
Leu122-Ser199-Arg426-Gly431	(1591)	AACAAAGGCCCTGGACAGATACTGGAACATC
Lys121-Val200-Asn425-Lys432	(1579)	AACAAAGGCCCTGGACAGATACTGGAACATC
Val120-Ile201-Ile424-Ala433	(1567)	AACAAAGGCCCTGGACAGATACTGGAACATC
Val120-Ile201B-Ile424-Ala433	(1567)	AACAAAGGCCCTGGACAGATACTGGAACATC
Consensus	(1621)	AACAAAGGCCCTGGACAGATACTGGAACATC
	1651	1680
Leu122-Ser199 Tryp427-Gly431	(1621)	ATGACCTGGATGGACGGGAGATC
Val127-Asn195-Arg426-Gly431	(1651)	ATGACCTGGATGGACGGGAGATC
Val120-Thr202-Ile424-Ala433	(1597)	ATGACCTGGATGGACGGGAGATC
Leu122-Ser199-Arg426-Lys432	(1621)	ATGACCTGGATGGACGGGAGATC
Leu122-Ser199-Arg426-Gly431	(1621)	ATGACCTGGATGGACGGGAGATC
Lys121-Val200-Asn425-Lys432	(1609)	ATGACCTGGATGGACGGGAGATC
Val120-Ile201-Ile424-Ala433	(1597)	ATGACCTGGATGGACGGGAGATC
Val120-Ile201B-Ile424-Ala433	(1597)	ATGACCTGGATGGACGGGAGATC
Consensus	(1651)	ATGACCTGGATGGACGGGAGATC
	1681	1710
Leu122-Ser199 Tryp427-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCTG
Val127-Asn195-Arg426-Gly431	(1681)	GACAACTACACCAACCTGATCTACACCTG
Val120-Thr202-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCTG
Leu122-Ser199-Arg426-Lys432	(1651)	GACAACTACACCAACCTGATCTACACCTG
Leu122-Ser199-Arg426-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCTG
Lys121-Val200-Asn425-Lys432	(1639)	GACAACTACACCAACCTGATCTACACCTG
Val120-Ile201-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCTG
Val120-Ile201B-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCTG
Consensus	(1681)	GACAACTACACCAACCTGATCTACACCTG

		1711	
Leu122-Ser199 Tryp427-Gly431	(1681)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Val127-Asn195-Arg426-Gly431	(1711)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Val120-Thr202-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Leu122-Ser199-Arg426-Lys432	(1681)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Leu122-Ser199-Arg426-Gly431	(1681)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Lys121-Val200-Asn425-Lys432	(1669)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Val120-Ile201-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Val120-Ile201B-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	
Consensus	(1711)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAAG	1770
	1741		
Leu122-Ser199 Tryp427-Gly431	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val127-Asn195-Arg426-Gly431	(1741)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val120-Thr202-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199-Arg426-Lys432	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199-Arg426-Gly431	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Lys121-Val200-Asn425-Lys432	(1699)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val120-Ile201-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val120-Ile201B-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Consensus	(1741)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	1800
	1771		
Leu122-Ser199 Tryp427-Gly431	(1741)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Val127-Asn195-Arg426-Gly431	(1771)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Val120-Thr202-Ile424-Ala433	(1717)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Leu122-Ser199-Arg426-Lys432	(1741)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Leu122-Ser199-Arg426-Gly431	(1741)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Lys121-Val200-Asn425-Lys432	(1729)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Val120-Ile201-Ile424-Ala433	(1717)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Val120-Ile201B-Ile424-Ala433	(1717)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	
Consensus	(1771)	TGGGCCAGCCCTGTGGAACCTGGTTGACATC	1830
	1801		
Leu122-Ser199 Tryp427-Gly431	(1771)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Val127-Asn195-Arg426-Gly431	(1801)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Val120-Thr202-Ile424-Ala433	(1747)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199-Arg426-Lys432	(1771)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199-Arg426-Gly431	(1771)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Lys121-Val200-Asn425-Lys432	(1759)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Val120-Ile201-Ile424-Ala433	(1747)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Val120-Ile201B-Ile424-Ala433	(1747)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	
Consensus	(1801)	AGCAAGCTGGCTGTGGTACATCAAGATCTTC	1860
	1831		
Leu122-Ser199 Tryp427-Gly431	(1801)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Val127-Asn195-Arg426-Gly431	(1831)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Val120-Thr202-Ile424-Ala433	(1777)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Leu122-Ser199-Arg426-Lys432	(1801)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Leu122-Ser199-Arg426-Gly431	(1801)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Lys121-Val200-Asn425-Lys432	(1789)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Val120-Ile201-Ile424-Ala433	(1777)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Val120-Ile201B-Ile424-Ala433	(1777)	ATCATGATCGTGGGGGGCTGGGGGGCTG	
Consensus	(1831)	ATCATGATCGTGGGGGGCTGGGGGGCTG	1890
	1861		
Leu122-Ser199 Tryp427-Gly431	(1831)	CGGATCGTGTCAACCGTGCTGAGCATCGTG	
Val127-Asn195-Arg426-Gly431	(1861)	CGGATCGTGTCAACCGTGCTGAGCATCGTG	
Val120-Thr202-Ile424-Ala433	(1807)	CGGATCGTGTCAACCGTGCTGAGCATCGTG	
Leu122-Ser199-Arg426-Lys432	(1831)	CGGATCGTGTCAACCGTGCTGAGCATCGTG	
Leu122-Ser199-Arg426-Gly431	(1831)	CGGATCGTGTCAACCGTGCTGAGCATCGTG	
Lys121-Val200-Asn425-Lys432	(1819)	CGGATCGTGTCAACCGTGCTGAGCATCGTG	

Val120-Ile201-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGTGAGCATCGTG
Val120-Ile201B-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGTGAGCATCGTG
Consensus	(1861)	CGCATCGTGTTCACCGTGTGAGCATCGTG
	1891	1920
Leu122-Ser199 Tryp427-Gly431	(1861)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Val127-Asn195-Arg426-Gly431	(1891)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Val120-Thr202-Ile424-Ala433	(1837)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Arg426-Lys432	(1861)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Arg426-Gly431	(1861)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Lys121-Val200-Asn425-Lys432	(1849)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Val120-Ile201-Ile424-Ala433	(1837)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Val120-Ile201B-Ile424-Ala433	(1837)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Consensus	(1891)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
	1921	1950
Leu122-Ser199 Tryp427-Gly431	(1891)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Val127-Asn195-Arg426-Gly431	(1921)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Val120-Thr202-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Leu122-Ser199-Arg426-Lys432	(1891)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Leu122-Ser199-Arg426-Gly431	(1891)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Lys121-Val200-Asn425-Lys432	(1879)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Val120-Ile201-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Val120-Ile201B-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
Consensus	(1921)	AGCTTCCAGACCCGCTTCGGGCCCCCGC
	1951	1980
Leu122-Ser199 Tryp427-Gly431	(1921)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Val127-Asn195-Arg426-Gly431	(1951)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Val120-Thr202-Ile424-Ala433	(1897)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Leu122-Ser199-Arg426-Lys432	(1921)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Leu122-Ser199-Arg426-Gly431	(1921)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Lys121-Val200-Asn425-Lys432	(1909)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Val120-Ile201-Ile424-Ala433	(1897)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Val120-Ile201B-Ile424-Ala433	(1897)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
Consensus	(1951)	GGCCCCGGACCCGCCCCGAGGCCATCGAGGAG
	1981	2010
Leu122-Ser199 Tryp427-Gly431	(1951)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Val127-Asn195-Arg426-Gly431	(1981)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Val120-Thr202-Ile424-Ala433	(1927)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Leu122-Ser199-Arg426-Lys432	(1951)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Leu122-Ser199-Arg426-Gly431	(1951)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Lys121-Val200-Asn425-Lys432	(1939)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Val120-Ile201-Ile424-Ala433	(1927)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Val120-Ile201B-Ile424-Ala433	(1927)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
Consensus	(1981)	GAGGGGGGGGAGGGGAGGGGAGGGGAGGGCAGC
	2011	2040
Leu122-Ser199 Tryp427-Gly431	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Val127-Asn195-Arg426-Gly431	(2011)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Val120-Thr202-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199-Arg426-Lys432	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199-Arg426-Gly431	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Lys121-Val200-Asn425-Lys432	(1969)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Val120-Ile201-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Val120-Ile201B-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Consensus	(2011)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
	2041	2070
Leu122-Ser199 Tryp427-Gly431	(2011)	ATCTGGGAGGGAGGGAGGGAGGGAGGGAGGG
Val127-Asn195-Arg426-Gly431	(2041)	ATCTGGGAGGGAGGGAGGGAGGGAGGGAGGG
Val120-Thr202-Ile424-Ala433	(1987)	ATCTGGGAGGGAGGGAGGGAGGGAGGGAGGG

Leu122-Ser199-Arg426-Lys432	(2011)	ATCTGGGACGACCTGCGGAGCCTGTGCGCTG
Leu122-Ser199-Arg426-Gly431	(2011)	ATCTGGGACGACCTGCGGAGCCTGTGCGCTG
Lys121-Val200-Asn425-Lys432	(1999)	ATCTGGGACGACCTGCGGAGCCTGTGCGCTG
Val120-Ile201-Ile424-Ala433	(1987)	ATCTGGGACGACCTGCGGAGCCTGTGCGCTG
Val120-Ile201B-Ile424-Ala433	(1987)	ATCTGGGACGACCTGCGGAGCCTGTGCGCTG
Consensus	(2041)	ATCTGGGACGACCTGCGGAGCCTGTGCGCTG
	2071	2100
Leu122-Ser199 Tryp427-Gly431	(2041)	TTCAGCTACCACCGCCTGCGGACCTGATC
Val127-Asn195-Arg426-Gly431	(2071)	TTCAGCTACCACCGCCTGCGGACCTGATC
Val120-Thr202-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGGACCTGATC
Leu122-Ser199-Arg426-Lys432	(2041)	TTCAGCTACCACCGCCTGCGGACCTGATC
Leu122-Ser199-Arg426-Gly431	(2041)	TTCAGCTACCACCGCCTGCGGACCTGATC
Lys121-Val200-Asn425-Lys432	(2029)	TTCAGCTACCACCGCCTGCGGACCTGATC
Val120-Ile201-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGGACCTGATC
Val120-Ile201B-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGGACCTGATC
Consensus	(2071)	TTCAGCTACCACCGCCTGCGGACCTGATC
	2101	2130
Leu122-Ser199 Tryp427-Gly431	(2071)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val127-Asn195-Arg426-Gly431	(2101)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val120-Thr202-Ile424-Ala433	(2047)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(2071)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(2071)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Lys121-Val200-Asn425-Lys432	(2059)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val120-Ile201-Ile424-Ala433	(2047)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(2047)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Consensus	(2101)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
	2131	2160
Leu122-Ser199 Tryp427-Gly431	(2101)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Val127-Asn195-Arg426-Gly431	(2131)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Val120-Thr202-Ile424-Ala433	(2077)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Arg426-Lys432	(2101)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Arg426-Gly431	(2101)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Lys121-Val200-Asn425-Lys432	(2089)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Val120-Ile201-Ile424-Ala433	(2077)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Val120-Ile201B-Ile424-Ala433	(2077)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
Consensus	(2131)	GGCCGCCCGGGCTGGGAGGCCCTGAAGTAC
	2161	2190
Leu122-Ser199 Tryp427-Gly431	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val127-Asn195-Arg426-Gly431	(2161)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Thr202-Ile424-Ala433	(2107)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Leu122-Ser199-Arg426-Lys432	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Leu122-Ser199-Arg426-Gly431	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Lys121-Val200-Asn425-Lys432	(2119)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Ile201-Ile424-Ala433	(2107)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Ile201B-Ile424-Ala433	(2107)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Consensus	(2161)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
	2191	2220
Leu122-Ser199 Tryp427-Gly431	(2161)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Val127-Asn195-Arg426-Gly431	(2191)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Val120-Thr202-Ile424-Ala433	(2137)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Leu122-Ser199-Arg426-Lys432	(2161)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Leu122-Ser199-Arg426-Gly431	(2161)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Lys121-Val200-Asn425-Lys432	(2149)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Val120-Ile201-Ile424-Ala433	(2137)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Val120-Ile201B-Ile424-Ala433	(2137)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
Consensus	(2191)	GAGCTGAAGAAGGGGGCGTGAACCTGTC
	2221	2250

FIG. 5M

Leu122-Ser199 Tryp427-Gly431	(2191)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Val127-Asn195-Arg426-Gly431	(2221)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Val120-Thr202-Ile424-Ala433	(2167)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199-Arg426-Lys432	(2191)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199-Arg426-Gly431	(2191)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Lys121-Val1200-Asn425-Lys432	(2179)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Val120-Ile201-Ile424-Ala433	(2167)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Val120-Ile201B-Ile424-Ala433	(2167)	GACGCCATGCCATCGCCGTGGCCGAGGGC
Consensus	(2221)	GACGCCATGCCATCGCCGTGGCCGAGGGC
		2251 2280
Leu122-Ser199 Tryp427-Gly431	(2221)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Val127-Asn195-Arg426-Gly431	(2251)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Val120-Thr202-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Leu122-Ser199-Arg426-Lys432	(2221)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Leu122-Ser199-Arg426-Gly431	(2221)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Lys121-Val1200-Asn425-Lys432	(2209)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Val120-Ile201-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Val120-Ile201B-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCAGCGC
Consensus	(2251)	ACCGACCGCATCATCGAGGTGGCCAGCGC
		2281 2310
Leu122-Ser199 Tryp427-Gly431	(2251)	ATCGGGGGGGGCGCTTCCATGCACATCCCCCGC
Val127-Asn195-Arg426-Gly431	(2281)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Val120-Thr202-Ile424-Ala433	(2227)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Leu122-Ser199-Arg426-Lys432	(2251)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Leu122-Ser199-Arg426-Gly431	(2251)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Lys121-Val1200-Asn425-Lys432	(2239)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Val120-Ile201-Ile424-Ala433	(2227)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Val120-Ile201B-Ile424-Ala433	(2227)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
Consensus	(2281)	ATCGGGGGGGCGCTTCCATGCACATCCCCCGC
		2311 2340
Leu122-Ser199 Tryp427-Gly431	(2281)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Val127-Asn195-Arg426-Gly431	(2311)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Val120-Thr202-Ile424-Ala433	(2257)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Leu122-Ser199-Arg426-Lys432	(2281)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Leu122-Ser199-Arg426-Gly431	(2281)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Lys121-Val1200-Asn425-Lys432	(2269)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Val120-Ile201-Ile424-Ala433	(2257)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Val120-Ile201B-Ile424-Ala433	(2257)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
Consensus	(2311)	CGCATCGCCCAAGGGGTTGGAGGGGGGGCTG
		2341 2352
Leu122-Ser199 Tryp427-Gly431	(2311)	CTGTAACCTCGAG
Val127-Asn195-Arg426-Gly431	(2341)	CTGTAACCTCGAG
Val120-Thr202-Ile424-Ala433	(2287)	CTGTAACCTCGAG
Leu122-Ser199-Arg426-Lys432	(2311)	CTGTAACCTCGAG
Leu122-Ser199-Arg426-Gly431	(2311)	CTGTAACCTCGAG
Lys121-Val1200-Asn425-Lys432	(2299)	CTGTAACCTCGAG
Val120-Ile201-Ile424-Ala433	(2287)	CTGTAACCTCGAG
Val120-Ile201B-Ile424-Ala433	(2287)	CTGTAACCTCGAG
Consensus	(2341)	CTGTAACCTCGAG

## SEQ ID NO:3 VAL120-ALA204

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCAGGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCGTG  
TGGAGGAGGCCACCACCAACCTGTTCTGCCAGCAGCCAAGGCCAACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCGCAGGCCAACGACCCCAACCCCCAGGAGATCGTGCT  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCTGTGGGACCAAGGCCCTGAAGGCCCTGCGTGGGCCGGCGCTGCCCAA  
GGTGAGCTTCGAGCCCATCCCCATCCACTACTGCCCCCCCGCCGGCTCGCCATCCTGAAGTG  
CAACGACAAGAAGTTCAACGGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCC  
ACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGAAACGGCAGCCTGGCGAGGAGGGC  
GTGGTGATCCGAGCGAGAACTTCACCGACAACGCCAACGACCATCATCGTGAGCTGAAGGA  
GAGCGTGGAGATCAACTGCACCCGCCAACAAACAACACCCGCAAGAGCATACCATCGGCC  
CCGGCCGCCCTTACGCCACCGCGACATCATCGGCCACATCCGCCAGGCCACTGCAACA  
TCAGCGCGAGAACAGTGGAAACAACACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTC  
GGCAACAAGACCATCGTGTCAAGCAGAGCAGCGGCCGGCACCCCGAGATCGTGATGCACAG  
CTTCAACTGCGCGGCAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAA  
CAACACCATCGGCCCAACAACACCAACGGCACCACATCCCTGCCCTGCCCATCAAGCAGA  
TCATCAACCGCTGGCAGGAGGTGGCAAGGCCATGTACGCCCCCCCATCCGCCAGGAGATC  
CGCTGCAGCAGCAACATCACCGCCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAA  
CACCAACCGAGATCTCCGCCCGCGCGGCGACATGCGCGACAACCTGGCGAGCGAGCTGT  
ACAAGTACAAGGTGGTAAGATCGAGCCCTGGCGTGGCCCCACCAAGGCCAAGGCCCGC  
GTGGTGAGCGAGAACGCGCCGTGACCCCTGGCGCCATGTTCTGGCTTCCTGGCGCC  
GCCGGCAGCACCATGGCGCCCGCAGCCTGACCCCTGACCGTGAGGCCGCCAGCTGCTGAG  
CGGCATCGTGCAGCAGCAGAACACACCTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTG  
AGCTGACCGTGTGGGCATCAAGCAGCTGCGAGGCCCGCTGCTGGAGCGCTACCTG  
AAGGACCAAGCAGCTGCTGGGCATCTGGGCTGAGCAGGCCAGCTGAGCTGACCCCTG  
GCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGAGCAAGCTGAGCTGACCCATGACCTGGA  
TGGAGTGGAGCGCGAGATCGACAACACCAACCTGATCTACACCTGATCTGACCCATGACCTGGA  
CAGAACCAAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGAGCAAGTGGCCAGCCTGT  
GGAACCTGGTCGACATCAGCAAGTGGCTGTGGTACATCAAAGATCTTCACTCATGATCGTGGCG  
GCCTGGTGGCCTGCGCATCGTGTACCGTGTGAGCATCGTAACCGCGTGCAGGCCAGGGCT  
ACAGCCCCCTGAGCTTCAGACCCGCTCCCCGCCCGCGGCCGCCAGCCGGAGGGCA  
TCGAGGAGGAGGGCGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG  
GCCCTGATCTGGACGACCTGCGCAGCCTGTCCTGTTCAAGTACCAACCGCCTGCGCGACCTG  
ATCCTGATCGCCGCCGATCGTGGAGCTGCTGGCCGCCGGCTGGAGGCCCTGAAGTAC  
TGGGGCAACCTGCTGAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTCGA  
CGCCATGCCATGCCGTGGCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCG  
GCCCGCCCTTCCCTGCACATCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAAC  
TCGAG

FIG. 6

## SEQ ID NO:4 VAL120-ILE201

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCGTG  
TGGAGGAGGCCACCACCAACCTGTTCTGCCAACGCCAGGCCAACGGCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCTCGTGCCTGCCAACGCCAACCCCCAAGGAGATCGTGCT  
GGAGAACCGTACCGAGAACTCAACATGTGAAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCGTGTGGGACCAAGAGCCTGAAGCCCTGCGTGGGCCATCACCCAGGCC  
CCCCAAGGTGAGCTCGAGCCCATCCCCATCCACTACTGCGCCCGGCCGCTCGCCATCCT  
GAAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCCTGCACCAACGTGAGCAGCAGCTGAGT  
GCACCCACGGCATCCGCCCCGTGGTGAAGCACCCAGCTGCTGAACGGCAGCCTGCCAG  
GAGGGCGTGGTATCCGAGCGAGAACTTCACCGACAACGCCAACGACATCGTGAGCAGCTGAG  
GAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAAACACCCGCAAGAGCATCACCA  
TCGGCCCCGGCGCGCCTCTACGCCACCGCGACATCATCGCGACATCCGCCAGGCC  
GCAACATCAGCGCGAGAAGGAAACACCCCTGAAGCAGATCGTGAACAGCTGAGCAGGCC  
CAGTCGGAACAAGACCATCGTGTCAAGCAGAGCAGCGGGCGGACCCGAGATCGTGT  
GCACAGCTTCAACTGCAGCGAGTTCTACTGCAACAGCAGCCAGCTGTTCAACAGCAC  
CTGGAACAAACACCATCGGCCCCAACAAACACCAACGGCACCATCACCTGCCGCATCA  
AGCAGATCATCAACCGCTGGCAGGAGGTGGCAAGGCCATGTACGCCCGCATCCGCC  
CAGATCCGCTGCAGCAGAACATCACCGCCCTGCTGCTGACCGCGACGGCGCAAGGAGAT  
CAGCAACACCAACCGAGATCTCCGCCCCGGCGGGGACATCGCGACAACTGGCGAGCG  
AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCAACCAAGGCAAG  
CGCCCGTGGTGCAGCGAGAAGCGCGCCGTGACCCCTGGCGCATGTTCTGGCTTCTG  
GGCGCCGCCGGCAGCACCATGGCGCCCGAGCGCTGACCCCTGACCGTGCAGGCC  
GCTGAGCGGCATCGTGCAGCAGCAGAACAAACCTGCTGCGGCCATCGAGGCCAGCAGCACC  
TGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGTGTGGCCGTGGAGCGC  
TACCTGAAGGACCAAGCAGCTGTGGCATCTGGGCTGACCGCAAGCTGATCTGCACCC  
CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAAGAGCCTGGACAGATCTGGAACAAACATGA  
CCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACCAACCTGATCTACACCC  
TGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGTGTGGCCGTGGAGCGC  
GAGAGCCAGAACCAAGCAGGAGAAGAACAGCAGCAGGAGCTGCTGGAGCTGGACA  
GCCTGTGGAACTGGTCGACATCAGCAAGTGGCTGTGTAACATCAAAAGATCTTC  
TGGCGGCCCTGGTGGCCCTGCGCATCGTGTTCACCGTGTGAGCATCGTGAACCGCG  
AGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTCCCCGCCCGCGGCC  
AGGGCATCGAGGAGGAGGGCGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGG  
CCTGCTGGCCCTGATCTGGAGCAGCTCGCGAGCGTGTGCTGTTCA  
CGACCTGATCCTGATCGCGCCCGCATCGTGGAGCTGCTGGCCGCCGGCTGGAGGCC  
GAAGTACTGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCC  
TGTCGACGCCATGCCATGCCGTGGCGAGGGCACCGACCGCATCGAGGTGGCC  
GCATCGGCCGCCCTCCTGCACATCCCCGCCCATCCGCCAGGGCTCGAGCGCGCC  
TGTAACCTCGAG

FIG. 7

## SEQ ID NO:5 VAL120-ILE201B

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCACTCTCG  
 TTTCGCCACGCCGTGGAGAACGCTGTGGGTGACCGTGTACTACGGCGTCCCCGTGAGAAGGAGCCA  
 CCACCAACCTGTTCTGCCAGCGACGCCAAGGCCCTACGACACCGAGGTGACAACGTGTGGGCCACCC  
 ACGCCTGCGTCCCCACCGACCCCAACCCCAAGGAGATCGTGTGGAGAACGTGACCGAGAACTTCAACA  
 TGTGGAAGAACAAACATGGTGGAGCAGATGCACGAGGACATCATCAGCTGTGGGACCAAGGCCCTGAAGC  
 CCTGCGTCCCCGGCATACCCAGGCCCTGCCCCAAGGTGAGCTTCAGGCCATCCCCATCCACTACTGC  
 CCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTCAACGGCAGGGGGCTGCAACCAACGT  
 GAGCACCGTGCAGTGACCCACGGCATCCGGCCCTGAGCAGGACCCAGCTGTGTGAAACGGCAGCCT  
 GGGCGAGGAGGGCTGGTGTACCGCAGCGAGAACCTCACCGACAACGCCAAGACATCATCGTGCAGCT  
 GAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAAACACCCGCAAGAGCATTACCGCC  
 CGGCCGCGCCTTCTACGCCACCGCGACATCATCGGCCACATCGCCAGGCCACTGCAACATCAGCG  
 GAGAAGTGGAACACACCCCTGAAGCAGATCGTGTGACCAAGCTGCAAGGCCAGTTCGGCAACAAAGAC  
 GTGTTCAAGCAGAGCAGCGCCGGGAGATCGTGTGACAGCTTCAACTCGGCCACACACCAAC  
 TTCTACTGCAACAGCACCGCAGCTTCAACAGCACCTGGAACAAACACCATCGGCCAACAAACACCAAC  
 GGCACCATCACCTGCCCTGCCATCAAGCAGATCATCAACCGCTGGCAGGAGTGGCAAGGCCATG  
 TACGCCGCCCATCGCCGGCAGATCGCTGCAGCACACATCACCGCCCTGCTGCTGACCCGGCAGC  
 GCGGCAAGGAGATCAGCAACACCCACCGAGATCTCCGCCCGGGGGGACATCGCGACAACCTGGC  
 GCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCCACCAAGGCCAAC  
 GCCCGTGGTGCAGCGAGAACGCGCCGTGACCCCTGGCGCATGTTCTGGCTTCTGGCGCC  
 CGGCAGCACCATGGCGCCCGAGCTGACCCCTGACCGTGCAGGCCAGCTGCTGAGCGG  
 GCAGCAGCAGAACAAACCTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCGTGTGGGG  
 CATCAAGCAGCTGCAAGGCCCGCTGGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGCAT  
 CTGGGGCTGAGCGGAAGCTGATCTGACCCACCGCCCTGGAACGCCAGCTGGAGCAACAGAG  
 CCTGGACAGATCTGGAACAACATGACCTGGATGGAGTGGAGGCCAGATCGACAAC  
 GATCTACACCCCTGATCGAGGAGAGCCAGAACAGCAGGAGAAGAACGAGCAGGAGCTGGAGCTGG  
 ACAAGTGGGCCAGCCTGGAACTGGTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTCAT  
 GATCGTGGGCCCTGGTGGGCCTGCGCATCGTGTGTTACCGTGTGAGCATCGTGAACCGCGTGC  
 GGCTACAGCCCCCTGAGCTTCAAGACCCGCTTCCCCGCCGGCCCCGACGCCCGAGGGCATCG  
 AGGAGGAGGGCGCGAGCGGACCGCAGCGCACGCCCTGGTGCACGCCCTGCTGGCCCTGATCT  
 GGGACGACCTGCGCAGCTGTGCTGTCAGCTACCAACCGCTGGGACCTGATCCTGATCGCC  
 CATCGTGGAGCTGCTGGCCGCCGCGCTGGAGGCCCTGAGACTGAGCTGGCAACCTGCTGCA  
 GATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTCAGGCCATGCCATCGCCGTGGCGAGGG  
 CGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCCCTTCTGACATCCCCGCCGATCCGCC  
 GGCTCGAGCGGCCCTGCTGTAACCGAGCGT

FIG. 8

## SEQ ID NO:6 LYS121-VAL200

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTCCCCGTG  
TGGAGGAGGCCACCAACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGT  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGGCCCTGCGTGAAGGCCCGTGTACACCCA  
GGCCTGCCCAAGGTGAGCTCGAGCCCATCCCCATCCACTACTGCCGCCCCGCCGCTCGC  
CATCCTGAAGTCAACGACAAGAAGTCAACGGCAGCGCCCCCTGACCAACGTGAGCACCG  
TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAACCCCCAGCTGCTGTAACGGCAGCCTGG  
CCGAGGAGGGCGTGGTGAATCCGCAAGAGAACTTACCGACAACGCCAAGACCATATCGT  
CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAAACAACACCCGCAAGAGCAT  
CACCATCGGCCCCGGCGCCCTTCAACGCCACCGCGACATCATCGCGACATCCGCCAGGC  
CCACTGCAACATCAGCGCGAGAACGAGATCGAACACACCCCTGAAGCAGATCGTACCGAAGCTGC  
AGGCCAGTTCGGCAACAAGACCATCGTGTCAAGCAGAGCAGCGGGGAGACCCGAGATC  
GTGATGCACAGTTCAACTGCAGCGGGAGTTCTACTGCAACAGCACCCAGCTGTTAAC  
AGCACCTGGAACAACACCATCGGCCCCAACAAACACCAACGGCACCATCACCTGCCCTGCCG  
CATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGCAAGGCCATGTACGCCGCGACGGCGAAG  
GAGATCAGCAACACCCACCGAGATCTCCGCCCGCGCGACATGCGGACAACCTGGCG  
CAGCGAGCTGTACAAGTACAAGGTGGTAAGATCGAGCCCCCTGGCGTGGCCCCAACCAAGG  
CCAAGCGCCGCGTGGTGCAGCGCGAGAACGCGCCCGTACCCCTGGCGCATGTTCTGGC  
TTCCTGGCGCCGCGCCGAGCACCATGGCGCCCGCAGCGTACCCCTGACCGTGCAGGCCCGC  
CAGCTGCTGAGCGGATCGTGCAGCAGCAGAACAAACCTGCTGCGGCCATCGAGGCCAGCA  
GCACCTGCTGAGCGTGTGGGCATCAAGCAGCTGAGGCCGCGTGTGGCGTGGCGTGG  
AGCGCTACCTGAAGGACCGAGCAGCTGCTGGCATCTGGGCTGAGCGGCAAGCTGATCTGC  
ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAAGATCTGAAACA  
CATGACCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACCAACCTGATCTACACCTGA  
TCGAGGAGAGCCAGAACCGAGCAGGAGAACAGAGCAGGAGCTGCTGGAGCTGGACAAGTG  
GCCAGCCTGTGGAACTGGTTGACATCAGCAAGTGGCTGTGGTACATCAAGATCTCAT  
GATCGTGGCGCCCTGGTGGCGCTGCGCATCGTGTTCACCGTGTGAGCATCGTAACCGCGT  
GCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGCGGCCCCGACCG  
CCCCGAGGGCATCGAGGAGGGCGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGC  
ACGGCCTGCTGCCCTGATCTGGGACGACCTGCGCAGCGTGTGCTGTTCAAGTACCGCC  
TGCAGCTGATCTGATGCCGCCATCGTGGAGCTGCTGGCTGGCGCGCTGGAG  
CCCTGAAGTACTGGGCAACCTGCTGCACTGGATCCAGGAGCTGAAGAACAGCGCCGTG  
AGCCTGTTGACGCCATGCCATGCCGTGGCCAGGGCACCGACCCGATCATCGAGGTGGC  
CAGCGCATCGGCCGCCCTTCTGCACATCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCC  
CTGCTGTAACTCGAGCGTGT

FIG. 9

## SEQ ID NO:7: LEU122-SER199

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCAGGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCTCGT  
TGGAGGAGGCCACCAACCACCTGTTCTGCCAGCGACGCCAACGGCTACGACACCGAGGT  
GCACAAACGTGTGGGCCACCCACGCCCTGCGTGCCTGCCACCGACCCAACCCCCAGGAGATCGTGC  
GGAGAACGTGACCGAGAACCTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGAAGCTGGCAACAGCGTGT  
CACCCAGGCCCTGCCCAAGGTGAGCTCGAGCCCCATCCCCATCCACTACTGCGCCCCCG  
CTTCGCCATCCTGAAGTGAACGACAAGAACGTTCAACGGCAGCGGGCCCTGCAACAAACGTGA  
GCACCGTGCAGTCAGCAGGCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGACCGC  
AGCCTGGCGAGGGAGGGCGTGGTGATCCGCAGCGAGAACCTCACCGACAACGCCAACGACAT  
CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAAACAACACCGCA  
AGAGCATACCATCGGCCCGGCCCTTCTACGCCACCGGCAGCATCATCGCGACATCC  
GCCAGGCCACTGCAACATCAGCGCGAGAACGAGTGGAACAAACACCCCTGAAGCAGATCGTGC  
AAGCTGCAGGCCAGTCGGAACAAGACCATCGTGTCAAGCAGAGCAGCGGGCGGCGACCC  
CGAGATCGTGTGACAGCTCAACTCGCGCGAGTTCTACTGCAACAGCACCCAGCT  
GTTCAACAGCACCTGGAACAAACACCATCGGCCCAACAAACACCAACGGCACCATCACCTGC  
CCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGCAAGGCCATGTACGCC  
CCCATCCGCCAGATCCGCTGCAGCAGCAACATCACCGCCCTGCTGCTGACCCGCGACGCC  
GGCAAGGAGATCAGCAACACCAACCGAGATCTCCGCCCGGCCGAGCATGCGCAGCAA  
CTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTAAGATCGAGCCCCCTGGCGTGGCCCC  
CCAAGGCCAAGCGCCGCGTGGTGAGCGCGAGAACGCGCCGTGACCCCTGGCGCCATGTT  
CTGGGCTTCCCTGGGCCGCCCGCAGCACCATGGCGCCCGCAGCCTGACCCCTGACCGTGCAG  
GCCCGCCAGCTGCTGAGCGGCATCGTCAGCAGCAGAACAAACCTGCTGCGGCCATCGAGGC  
CCAGCAGCACCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAAGGCCGCGTGTGG  
CCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTGGCATTGGGCTGCAAGGGCAAGCTG  
ATCTGCACCAACGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACAGATCTG  
GAACAAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACCAACCTGATCTACA  
CCCTGATCGAGGAGAGCCAGAACAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGA  
CAAGTGGCCAGCCTGTGGAACCTGGTCAACATCAGCAAGTGGCTGTGGTACATCAAGATCTT  
CATCATGATCGTGGCGGCCCTGGTGGCCTGCCATCGTGTTCACCGTGTGAGCATCGTGA  
CCCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCAGACCCGCTTCCCGCCCGCGGCC  
CGACCGCCCCGAGGGCATCGAGGAGGGCGCGAGCGCAGCCGACCGCAGCAGCCCC  
CTGGTGCACGCCCTGCTGGCCCTGATCTGGACGACCTGCGCAGCCTGCTGCTTCAAGCTAC  
CACCGCCTGCGCAGCTGATCTGATCGCCGCCGCATCGTGGAGCTGCTGGCGCCGCC  
TGGGAGGCCCTGAAGTACTGGGCAACCTGCTGCAAGTACTGGATCCAGGAGCTGAAGAACAG  
CGCCGTGAGCCTGTTGACGCCATGCCATGCCGTGGCCAGGGCACCAGCAGCATCATCGA  
GGTGGCCAGCGCATCGGCCCTTCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGA  
GCGGCCCTGCTGTAACTCGAGCGTGT

FIG. 10

## SEQ ID NO:8 VAL120-THR202

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGGCCGTGGAGAACGCTGTGGGTACCGTGTACTACGGCGTCCCCGTG  
TGGAGGAGGCCACCAACCACCTGTTCTGCGCCAGCGACGCCAACGCCTACGACACCGAGGT  
GCACAACGTGTGGCCACCCACGCCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGT  
GGAGAACGTGACCGAGAACITCAACATGTGGAAGAACACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGACCAAGAGCCTGAAGGCCCTGCGTGGCGGCCACCCAGGCC  
CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGQCACCT  
GAAGTGCACAGACAAGAACGTTCAACGGCAGCGGCCCTGCAACCAACGTGAGCACCGTGCAGT  
GCACCCACGGCATCCGCCCGTGGTAGCAGCACCCAGCTGCTGTAACGGCAGCCTGGCCAG  
GAGGGCGTGGTATCCGCAGCGAGAACITCACCGACAACGCCAACGACATCATGTGCAGCT  
GAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAACACACCCGCAAGAGCATCACCA  
TCGGCCCCGGCGCCCTCTACGCCACCGCGACATCATGGCGACATCCGCCAGGCCACT  
GCAACATCAGCGCGAGAACGTTGAACAAACACCCCTGAAGCAGATCGTACCGAAGCTGCAGGC  
CAGTTCGGCAACAAGACCATCGTGTCAAGCAGAGCAGCGGCCGACCCGAGATCGTGT  
GCACAGCTCAACTCGGGCGCGAGTTCTACTGCAACAGCACCCAGCTGTTAACACGAC  
CTGGAACAACACCATCGGCCCCAACAAACACCAACGGCACATCACCTGCCCGCATCA  
AGCAGATCATCAACCGCTGGCAGGAGGTGGCAAGGCCATGTACGCCCTGGCGCC  
CAGATCCGCTGCAGCAGCAACATCACCGCCCTGCTGCTGACCCCGCAGGCCAGGAGAT  
CAGCAACACCCACCAGAGATCTTCGCCCCGGCGCGACATCGCGCACAACCTGGCGCAGCG  
AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCACCAAGGCCAG  
CGCCGCGTGGTGCAGCGCGAGAACGCGCCGTGACCCCTGGCGCCATGTTCTGGCTTCCT  
GGCGCCGCCGCAGCACCATGGCGCCCGCAGCCTGACCCCTGACCGTGCAGGCCAGCT  
GCTGAGCGGCATCGTCAGCAGCAGAACAAACCTGCTGCGGCCATCGAGGCCAGCAGCACC  
TGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGTGTGGCGTGGAGCGC  
TACCTGAAGGACCAGCAGCTGCTGGCATCTGGGCTGCAAGCAGCTGATCTGCACCCAC  
CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAAGCTGATCTGGAAC  
CCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACCAACCTGATCTACACCC  
TATCGAGGAGGCCAGAACCGCAGGAGAACAGCAGCAGGAGCTGCTGGAGCTGGACAAG  
GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTTACATCAAGATCTTACATGATCG  
TGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCACCGTGTGAGCATCGTGAACCGCG  
AGGGCTACAGCCCCCTGAGCTTCAGACCCGCTTCCCCGCCCGCGGCCGACCGCCCG  
AGGGCATCGAGGAGGAGGGCGCGAGCGCGACCCGCAAGCAGGCCCTGGTGCACGG  
CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGCTGCTGTTAGCTACCAACCC  
CGACCTGATCCTGATCGCCGCCGCATCGTGGAGCTGCTGGGCCCGCGCTGGAGGCC  
GAAGTACTGGGCAACCTGCTGAGTACTGGATCCAGGAAGCTGAAGAACAGCGCC  
TGGTGCAGCCATGCCATGCCGTGGCGAGGGCACCGACCGCATCATCGAGGTGGCC  
GCATCGGCCGCCCTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC  
TGAACTCGAG

FIG. 11

## SEQ ID NO:9 TRP427-GLY431

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGGCCGGTGGAGAAGCTGTTGGTACCGTGACTACGGCGTGCCGTG  
TGGAGGAGGCCACCAACACCCCTGTTCTGCCAGCGACGCCAAGGCCAACGGAGGT  
GCACAACCGTGTGGGCCACCCACGCCCTCGTGTGCCACCGACCCCAACCCCCAGGAGATCGTGC  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCCTGAAGGCCCTGCGTGAAGCTGACCCCCCTGTGCGT  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAACCAAGAGCAGCAACTGGAAAGGAGAT  
GGACCGCGCGAGATCAAGAACGACTGCAGCTCAAGGTGACCAAGCAGCATCCGAACAAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGCCATCGACAAACGACAACACCAGC  
TACAAGCTGATCAACTGCAACACCAAGCGTGATCACCCAGGCCCTGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC  
CCGTGGTGAGCACCAGCTGCTGAACGGCAGCCTGGCGAGGAGGGCTGGTATCCGC  
AGCGAGAACCTCACCGACAACGCCAAGACCATCATCGTGAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCAACAACACACCCGCAAGAGCATACCATGGCCCCGGCGCCT  
TCTACGCCACCGGCGACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGCGAG  
AAAGTGGAAACAACACCTGAAGCAGATCGTACCAAGCTGCAAGGCCAGTTCGGCAACAAGAC  
CATCGTGTCAAGCAGAGCAGCGCGAGATCGTGATGCACAGCTCAACTCG  
GCGCGAGTTCTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG  
GCCCAACAACACCAACGGCACCATCACCCCGCCATCGCGCATCAAGCAGATCATCAACCGCT  
GGGGCGGCAAGGCCATGTACGCCCCCCCCATCCGCGGAGATCCGCTGCAGCAGCAACATC  
ACCGCCTGCTGCTGACCCCGACGGCGCAAGGAGATCAGCAACACCAACCGAGATCTTCCG  
CCCCGGCGCGCGACATGCGCGACAACACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGA  
AGATCGAGCCCCCTGGCGTGGCCCCACCAAGGCCAAGCGCCGGTGGTGCAGCGAGAAG  
CGCGCCGTGACCCCTGGCGCATGTTCTGGCTTCTGGCGCCGGCAGCACCATTGGC  
GCCCGCAGCCTGACCCGTGACCGTGCAAGGCCAGCTGTCAGCGGATCGCAGCAGCA  
GAACAAACCTGCTGCGCGCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGCA  
TCAAGCAGCTGCAGGCCCGCTGCTGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG  
GGCATCTGGGCTGCTGAGCGCAAGCTGATCTGCAACCACGCCGTGCCCTGGAACGCCAGCTG  
GAGCAACAAGAGCCTGGACCAAGATCTGGAACACATGACCTGGATGGAGTGGAGCGCGAG  
ATCGACAACTACACCAACCTGATCTACACCCGTGAGGAGAGGCCAGAACCAAGCAGGAGAA  
GAACGAGCAGGAGCTGGAGCTGGACAAGTGGGCCAGCCTGTGAACTGGTTCGACATCA  
GCAAGTGGCTGTGGTACATCAAGATCTTCACTATGATCGTGGCGGCCCTGGTGGGCTGCGCA  
TCGTGTTCACCGTGTGAGCATCGTAACCGCGTGCAGGCCAGGGCTACAGCCCCCTGAGCTCC  
AGACCCGCTTCCCCGCCCGCCCGAGGGCATCGAGGAGGGAGGCCAGCG  
GAGCGCAGCCGCACCGCAGCAGGCCCTGGTGCAGGCCCTGCTGCCCTGATCTGGGACGA  
CCTGCGCAGCCTGCTGCTGTCAGCTACCAACGCCCTGCGCGACCTGATCCTGATCGCCGCCG  
CATCGTGGAGCTGCTGGGCCCGCCGGTGGAGGCCCTGAAGTACTGGGCAACCTGCTGC  
AGTACTGGATCCAGAGCTGAAGAACAGGCCGTGAGCCTGTTGACGCCATGCCATCGCC  
GTGGCGAGGGCACCGACCGCATCGAGGTGGCCAGCGCATCGGCCCTGCTGTAACCTCGAG  
CATCCCCCGCCGATCCGCCAGGGCTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 12

## SEQ ID NO:10 ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGGCCGGTGGAGAACGCTGTGGGTGACCGTGACTACGGCGTGCCGTG  
TGGAGGAGGCCACCAACCACCTGTTCTGCCAGCAGCCAAGGCCAACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCTCGTGCCTGCCACCGACCCCAACCCCCAGGAGATCGTGT  
GGAGAACGTGACCGAGAACCTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCCTGAAGGCCCTGCGTGAAGCTGACCCCCCTGCGTG  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAACCCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGCGAGATCAAGAACGACTGCAGCTCAAGGTGACCAACCGAGCATCGCAACAAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGCCCATCGACAACGACAACACCAGC  
TACAAGCTGATCAACTGCAACACCAGCGTGTACCCAGGCCCTGCCATCTGAAGTGCAACGACAAGAA  
GCCCATCCCCATCCACTACTGCGCCCGGCCGCTGCCATCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTCAGTGCACCCACGGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGAACGGCAGGCCCTGGCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAACCTCACCGACAACGCCAGACCATCATCGTGAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCAACAAACACCCGCAAGAGCATCACCATCGGCCCGGCCCT  
TCTACGCCACCGGCACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGCGAG  
AAGTGGAAACAACACCCCTGAAGCAGATCGTACCGAACAGCTGCAGGCCAGTCCGCAACAAGAC  
CATCGTGTCAAGCAGAGCAGCGGCCGACCCAGAGATCGTGATGCACAGCTTCAACTGCG  
GCCCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAAACAACACCATCG  
GCCCAACAAACACCAACGGCACCATCACCTGCCCTGCCATCAAGCAGATCATCACCGC  
GGCGCGGCCAGGCCATGTACGCCCCCCCCTGCCCGCCAGATCCGCTGCAGCAGCAACAT  
CACCGGCCCTGCTGCTGACCCCGACGGCGCAAGGAGATCAGCAACACCAACCGAGATCTCC  
GCCCGGCCGCCGACATCGCGACAACTGGCGAGCGAGCTGTACAAGTACAAGTGGTG  
AAGATCGAGCCCCCTGGCGTGGCCCCCAACAGGCCAGGCCCGCTGGTGAGCGCAGAA  
GCCCGCCGTGACCCCTGGCGCCATGTTCTGGCTTCTGGCGCCGCCAGCACCATTGG  
CGCCCGCAGCCTGACCCCTGACCGTGAGGCCAGCTGAGGCCATCGTGAGCAGC  
AGAACAAACTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGC  
ATCAAGCAGCTGCAAGGCCCGCTGCTGCCGTGGAGCGTACCTGAAGGACCAGCAGCTGCT  
GGCATCTGGGCTGCAAGCGCAAGCTGATCTGCACCAACGCCGTGCCCTGGAAACGCCAGCT  
GGAGCAACAAGAGCCTGGACCAAGATCTGAACAAACATGACCTGGATGGAGTGGAGGCCGA  
GATCGACAACACACCAACCTGATCTACACCCCTGATCGAGGAGAGGCCAGAACAGCAGGAGA  
AGAACGAGCAGGAGCTGCTGGAGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC  
AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGCGGCCCTGGTGGGCCCTGCG  
ATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGCAGGCCAGGGCTACAGCCCCCTGAGCTTC  
CACAGGCCCTCCCCGCCCGGCCGAGGGCATCGAGGAGGAGGGCG  
CGAGCGCAGCGCACCGCAGCAGGCCCTGGTCACGGCTGCTGGCCCTGATCTGGAGC  
ACCTGCGCAGCCTGCTGCTGAGCTACCGCCTGCCGAGCTGATCCTGATCGCCGCC  
GCATCGTGGAGCTGCTGGGCCGCCGCGCTGGAGGCCCTGAAGTACTGGGCAACCTGCTG  
CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATGCCATCGC  
CGTGGCCGAGGGACCCGACCGCATCGAGGTTGCCAGCGCATGGCCGCCCTCCTGC  
ACATCCCCCGCCGACATCGCCAGGGCTCGAGCGCCCTGCTGTAACTCGAG

FIG. 13

## SEQ ID NO:11 ARG426-GLY431B

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTCGTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAGGAGGCCACCACCAACCTGTCGCCAGCGACGCCAAGGCCCTACGACACCGAGGT  
GCACAAACGTGTGGGCCACCCACGCCCTCGTGCCTGCCACCGACCCCAACCCCAAGGAGATCGTGC  
GGAGAACCGTGAACGAGAACTCAACATGTGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGAGCCTGAAGGCCCTGCGTGAAGCTGACCCCCCTGCGTG  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAACCCAAGGCAGCAACTGGAAGGAGAT  
GGACCGCGGCAGAGATCAAGAACTGCAGCTCAAGGTGACCACCGATCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTCTACAAGCTGGACGTGGTGCCTCGACAAACGACAACACCCAGC  
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCCTGCCATCCTGAAGTGCAACGACAAGAA  
GCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTCGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGACCAACGTGAGCACCGTGCAGTGACCCACGGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCGAGGAGGGCGTGGTGAATCCGC  
AGCGAGAACCTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCCAACAAACAACACCCGCAAGAGCATCACCATCGGCCCGCGCCT  
TCTACGCCACCGGCGACATCATCGGCCACATCGGCCAGGGCCACTGCAACATCAGCGGCAG  
AAAGTGAACACACCCCTGAAGCAGATCGTGAACAGCTGAGGCCAGTCCGCAACAAAGAC  
CATCGTGTCAAGCAGAGCAGCGCGCGACCCCGAGATCGTGTGACAGCTTCACAGCTCAACTCG  
GCGCGAGTTCTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAAACCCATCG  
GCCCAACAAACACCAACGGCACCATCACCTGCCCTGCCATCAAGCAGATCATCAACCGC  
GGCAGCGCAAGGCCATGTACCCCCCCCCATCCGCGCCAGATCCGCTGCGAGCAACAT  
CACCGGCCCTGCTGCTGACCCCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTCC  
GCCCGGCCGGCGGCGACATCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTG  
AAAGATCGAGCCCCCTGGCGTGGCCCCCACCAAGGCCAAGGCCAGCGCGCTGGTGCAGCGAGAA  
GCGCGCCGTGACCCCTGGCGCCATGTTCTGGCTCTGGCGCCGCCAGCACCATGGG  
CGCCCGCAGCCTGACCCCTGACCGTGAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC  
AGAACAAACCTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGCGAGCTGACCGTGTGGGC  
ATCAAGCAGCTGCAAGGCCCGTGTGGCGCTGGAGCGCTACCTGAAGGACCAAGCAGCTGCT  
GGGCATCTGGGCTGAGCGCAAGCTGATCTGCAACCACCGCCGTGCCCTGGAACGCCAGCT  
GGAGCAACAAGAGCCTGGACCAGATCTGGAACAAACATGACCTGGATGGAGTGGAGCGCGA  
GATCGACAACACTACACCAACCTGATCTACACCTGATCGAGGAGAGGCCAGAACCCAGCAGGAGA  
AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGCCAGCCTGTGGAACCTGGTCAACTG  
AGCAAGTGGCTGTGGTACATCAAGATCTCATGATCGTGGCGGCCCTGGTGGGCCTGCGC  
ATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGCAGCCAGGGCTACAGCCCCCTGAGCTTC  
CAGACCCGCTTCCCCGCCCGCCCCCGCGGCCAGCGCCCCGAGGGCATCGAGGAGGAGGGCG  
CGAGCGCAGCGCACCGCAGCAGCAGCCCCCTGGTGCAGGCCCTGCTGGCCCTGATCTGGAGC  
ACCTGCGCAGCCTGCGCTGTCAGCTACCAACGCCCTGCGCGACCTGATCTGATGCCGCC  
GCATCGTGGAGCTGCTGGCGCCGCGCTGGAGGGCCCTGAAGTACTGGGCAACCTGCTG  
CACTGAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC  
CGTGGCGAGGGCACCGACCCATCGAGGTGGCCAGCGCATCGGCCCTGCGCCTTCC  
ACATCCCCGCCGACATCCGCCAGGGCTCGAGCGCGCCCTGCTGTAACCGAG

FIG. 14

## SEQ ID NO:12 ARG426-LYS432

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCAGGCCGTGGAGAACGCTGTGGGTGACCGTGTACTACGGCGTGCCGTG  
TGGAGGAGGCCACCAACACCCCTGTTCTGCCAGCAGCCAAGGCCAACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCTCGTGCCTGCCACCGACCCCAACCCCCAGGAGATCGTGCT  
GGAGAACGTGACCGAGAACCTCAACATGTGGAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCCTGAAGGCCCTGCCGTGAAGCTGACCCCCCTGTGCGTG  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGGCAGAGATCAAGAACCTGAGCAGCTTCAAGGTGACCAACGCCATCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGCCATCGACAACGACAACACCCAGC  
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCCTGCCATCCTGAAGTGCAACGACAAGAA  
GCCCATCCCCATCCACTACTGCGCCCCCGCCGCTCGCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGTAACGGCAGCCTGGCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAACCTCACCGACAACGCCAACGACCATCATCGTGAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCAACAACACACCCGCAAGAGCATCACCATGGCCCCGGCGCCT  
TCTACGCCACCGGCACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGGAG  
AAAGTGAACAAACACCCCTGAAGCAGATCGTACCGAACAGCTGCAGGCCAGTTCGGCAACAAGAC  
CATCGTGTCAAGCAGAGCAGCGCGCGAACCCGAGATCGTATGCAAGCTTCAACTGCG  
GCCCGAGGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAAACACCATCG  
GCCCAACAAACACCGCACCATCACCTGCCCTGCCCATCAAGCAGATCATCAACCGC  
GGCGCAACAAGGCCATGTACGCCCCCCCACATCCCGGCCAGATCCGCTGCAAGCAGAACAT  
CACCGCCTGCTGCTGACCCGCAACGGCAAGGAGATCAGCAACACCACCGAGATCTTCC  
GCCCGGCCGGCGCACATCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG  
AAAGATCGAGCCCCCTGGCGTGGCCCCACCAAGGCCAACGCCCGTGGTGAGCGCAGAA  
GCGCGCCGTGACCCCTGGCGCCATGTTCTGGCTTCTGGCGCCGGCAGCACCATGGG  
CGCCCGCAGCCTGACCCCTGACCGTGAGGCCAGCTGCTGAGCGGATCGCAGCAGCAGC  
AGAACAAACCTGCTGCCGCCATCGAGGCCAGCAGCACCTCTGCAAGCTGACCGTGTGGGG  
ATCAAGCAGCTGCAAGGCCCGTGTGGCGTGGAGCGTACCTGAAGGAGCAGCAGCTGCT  
GGGCATCTGGGGCTGCAAGCGCAAGCTGATCTGACCCACCGCCGTGCCCTGGAACGCCAGCT  
GGAGCAACAAGAGCCTGGACCAAGATCTGGAACAAACATGACCTGGATGGAGTGGAGCGCAG  
GATCGACAACACCAACCTGATCTACCCCTGATCGAGGGAGGCCAGAACCCAGCAGGAGA  
AGAACGAGCAGGAGCTGCTGGAGCGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC  
AGCAAGTGGCTGTGTACATCAAGATCTTCACTGATCGTGGCGGCTGGTGGGCCCTGCGC  
ATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTT  
CAGACCCGCTTCCCCGCCCGGCCGAGGCCGAGGCCGAGGGCATCGAGGAGGAGGGCG  
CGAGCGCACCGCACCGCAGCAGGCCCTGGTGCACGCCCTGCCGACCTGATCTGGACG  
ACCTGCGCAGCCTGCGCTGTCAGTACACCCGCTGCCGACCTGATCTGATGCCGCC  
GCATCGTGGAGCTGCTGGGCCGCCGCGCTGGAGGCCCTGAAGTACTGGGCAACCTGCTG  
CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATGCCATCGC  
CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCCCTGCTG  
ACATCCCCCGCCGACATCCGCCAGGGCTTCGAGCGCCCTGCTGTAACCTCGAG

FIG. 15

## SEQ ID NO:13 ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGCGCCGTGGAGAAAGCTGTGGGTGACCGTGACTACGGCGTGCCGTG  
TGGAGGAGGCCACCAACCCCTGTTCTGCGCCAGCAGCCAAGGCCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCCAACCCCCAGGGAGATCGTGCT  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCCTGAAGGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGCGAGATCAAGAACGACTGCAGCTTCAAGGTGACCAACAGCATCGCAACAAAGATGC  
AGAAGGAGTACGCCCTGTTACAAGCTGGACGCTGGTGCCTGACAAACGACAACACCAGC  
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCCTGCCCCAACGGTGAAGCTTCA  
GCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTCGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGCCCCCTGACCAACGTGAGCACCGTGAGTCAGTGCACCCACGGCATCCGCC  
CCGTGGTGAAGCACCAAGCTGCTGTAACGGCAGGCCCTGGCGAGGAGGGCGTGGTGAATCCGC  
AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCAACAAACAACACCCGCAAGAGCATACCATCGGCCCCGGCGCCT  
TCTACGCCACCGGGACATCATCGGCACATCCGCCAGGCCACTGCAACATCAGCGGGAG  
AAAGTGAACAAACACCCCTGAAGCAGATCGTACCAAGCTGCAGGCCAGTTCGGCAACAAAGAC  
CATCGTGTCAAGCAGAGCAGCGGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCG  
GCCCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAAACACCATCG  
GCCCAACAAACACCAACGGCACCATCACCTGCCCTGCCCATCAAGCAGATCATCAACGCC  
CCAAGGCCATGTACCCCCCCCCATCCGCGCCAGATCCGCTGCAAGCAGCAACATCACC  
TGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCCAGAGATCTTCCGCCCCGC  
GGCGCGACATCGCCGACAACCTGGCGAGCGAGCTGACCAAGTCAAGGTGGTAAGATCGA  
GCCCTGGCGTGGCCCCACCAAGCCAAGCGCCGCGTGGTGCAGCGAGAACGCGCG  
TGACCCCTGGCGCCATGTTCTGGCTTCTGGCGCCGCGCAGCTGAGCGGCATCGTGAGCAG  
GCCCTGACCCCTGACCGTGAGGCCAGCTGCTGAGCGGCATCGTGAGCAGCAGAAC  
CTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCGTGTGGGCATCAAGCA  
GCTGAGGCCCGCGTGTGGCGCTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCT  
GGGGCTGCAAGCTGATCTGACCAACCGCCGTGCCCTGGAAACGCCAGCTGGAGCAAC  
AAAGAGCCTGGACCAAGATCTGGAACAAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAA  
CTACACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACGAGC  
AGGAGCTGCTGGAGCTGGACAAGTGGCCAGCCTGTGGAACCTGGTGCACATCGCAAGTGG  
CTGTGGTACATCAAGATCTTCACTCATGATCGTGGCGGCCTGGTGGGCCTGCGCATCGTGT  
ACCGTGTGAGCATCGTAACCGCGTGCAGGCCAGGGCTACAGCCCCCTGAGCTTCAAGACCC  
TTCCCCGCCCGGGCCCCGAGGGCATCGAGGAGGGAGGGCGAGCGCG  
CCCGGACCGCAGCAGCCCCCTGGTGCACGGCTGCTGGCCCTGATCTGGAGCACCTGCGCAG  
CCTGTGCCCTGTTCAAGCTACCAACCGCCCTGCGCAGCCTGATCCTGATCGCCGCCCGCATCGTGG  
GCTGCTGGCCGCCGGCTGGAGGCCCTGAAGTACTGGGCAACCTGCTGAGCTG  
TCCAGGAGCTGAAGAACAGGCCGTGAGCCTGTTGACGCCATGCCATGCCGTGGCGAG  
GGCACCGACCGCATCGAGGTGGCCAGCGCATCGGCCGCGCTTCTGACATCCCCCGC  
CGCATCCGCCAGGGCTTCGAGCGGCCCTGCTGTAACTCGAG

FIG. 16

## SEQ ID NO:14 ILE424-ALA433

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCCAGGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCGTG  
TGGAGGAGGCCACCAACCCCTGTTCTGCCAGCGACGCCAACGCCAACGACACCCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCGCGTGCCTGCCAACGACCCAACCCAGGAGATCGTGC  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCGTGTGGGACCAAGGCCCTGAAGGCCAACACCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGCGAGATCAAGAACGACCGTCAAGGTGACCACCAGCATCCGCAACAAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCTGCACAAACGACAAACCCAGC  
TACAAGCTGATCAACTGCAACACCCAGCGTGTACCCAGGCCCTGCCATCCTGAAGTGCAACGACAAGAA  
GCCCATCCCCATCCACTACTCGGCCCGCCGTTGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGTGACCCACGGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAACCTCACCGACAACGCCAACGACCATATCGTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCAACAAACAACACCCGCAAGAGCATCACCATGGCCCGCCGCCT  
TCTACGCCACCGCGACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGCGAG  
AAAGTGGAAACAACACCCCTGAAGCAGATCGTACCAAGCTGCAAGGCCAGTTGGCAACAAAGAC  
CATCGTGTCAAGCAGAGCAGCGCGCCGACCCGAGATCGTGTACAGCTTCAACTGCG  
GCCCGAGTTCTTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAAACCCATCG  
GCCCAACAAACACCAACGGCACCACCATCACCCTGCCCTGCCCATCAAGCAGATCATCGCGGC  
GCCATGTACGCCCGCCCATCCGCGCCAGATCCGCTGCAAGCAGCAACATCACCGCCCTGCTG  
CTGACCCCGCAGGGCGCAAGGAGATCAGCAACACCACCGAGATCTCCGCCCGGGCG  
CGACATGCGCAGCAACTGGCGAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCC  
TGGCGTGGCCCCACCAAGGCAAGCGCCGCGTGGTGCAGCGAGAAGCGCGCCGTGACC  
CTGGCGCCATGTTCTGGGCTTCTGGCGCCCGGCGACCATGGCGCCCGAGCCTG  
ACCTGACCGTGCAGGCCAGCTGCTGAGCGCATCGTGCAGCTGACCGTGTGGGCACTAAGCAGCTG  
GCCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGCACTTCCAGACCGCTTCC  
AGGCCCGCGTGTGGCGTGGAGCGTACCTGAAGGACCAAGCAGCTGCTGGGCACTTGGG  
TGCAGCGGCAAGCTGATCTGACCAACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG  
CCTGGACCAAGATCTGAACAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAACCTACA  
CCAACCTGATCTACACCCGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGA  
GCTGCTGGAGCTGGACAAGTGGCCAGCCTGTGGAACCTGGTGCACATCAGCAAGTGGCTGT  
GGTACATCAAGATCTTCACTGATCGTGGCGCCCTGGTGGGCGATCGCAGTGTGTTACCG  
TGCTGAGCATCGTGAACCGCGTGCAGGCCAGGGCTACAGCCCTGAGCTTCCAGACCGCTTCC  
CCGCCCGCCGCGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGAGCGCAGCG  
GACCGCAGCAGCCCCCTGGTGCAGGCCCTGCTGGCCCTGATCTGGAGCGACCTGCGCAGCG  
TGCCTGTTCACTGACCGCCCTGCGCGACCTGATCTGATCGCCGCCGCACTGAGCTG  
CTGGGCCGCCCGCGCTGGAGGCCCTGGAAGTACTGGGCAACCTGCTGCAGTACTGGATCCA  
GGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATGCCATGCCGTGGCGAGGGCA  
CCGACCGCAGTACGAGGTGGCCAGCGCATGCCATGCCGTGGCGCTTCCCTGCACATCCCCGCC  
TCCGCCAGGGCTTGCAGCGCCCTGCTGTAACCGAG

FIG. 17

## SEQ ID NO:15 ILE423-MET434

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCCAGCGCCGTGGAGAACGCTGTGGGTGACCGTGACTACGGCGTGCCTG  
TGGAGGAGGCCACCAACCCCTGTCGCCCCAGCAGCCAAAGGCCTACGACACCCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCTGCGTGCCTGCCCCACCGACCCAAACCCCAAGGAGATCGT  
GGAGAACGTGACCGAGAACCTCAACATGTGGAGAACAAACATGGTGGAGCAGATGCA  
GACATCATCAGCCTGTCGACCCAGGCTGAAGCAGCCACCAAGCAGCAACTGGAGGAGAT  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAAGCAGCAACTGGAGGAGAT  
GGACCGCGCGAGATCAAGAACGCTGAGCTCAAGGTGACCCACCGATCCGAACAAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCTGACAAACGACA  
TACAAGCTGATCAACTGCAACACCAGCGTGTACCCAGGCCCTGCCCAAGGTGAGCTCGA  
GCCCATCCCCATCCACTACTGCGCCCCGCCGGCTCGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGACCAACGTGAGCAGCGTGCAGTGACCCACGGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGTAACGGCAGCCTGCCGAGGAGGGCGTGGTATCCGC  
AGCGAGAACCTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCCAACAAACAACCCCGCAAGAGCATCACCATCGGCCCGCGCCT  
TCTACGCCACCGCGACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGCGAG  
AAGTGGAAACAACCCCTGAAGCAGATCGTACCAAGCTGCAGGCCAGTTGGCAACAAGAC  
CATCGTGTCAAGCAGAGCAGCGCGCAGCCGAGATCGTGTACAGCTTCAACTCG  
GCCCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACCCATCG  
GCCCAACAACCCAACGGCACCATCACCCCTGCCCTGCCCATCAAGCAGATCGCGCGATG  
TACGCCCTCCATCCCGGCCAGATCCGCTGCAAGCAGCAACATCACCGCCCTGCTGCTGACC  
CGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTCCGCCCGCGCGACAT  
GCGCAGCAACTGGCGAGCGAGCTGACAAGTACAAGGTGGTGAAGATCGAGGCCCTGGCG  
TGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGAGAACGCGCGAGCTGCCGACCTG  
GCCATGTTCTGGCTTCTGGCGCCCGCAGCACCATGGGCCCGCAGCTGCCGACCTG  
ACCGTGCAGGCCGCGCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACACCTGCTGCGCG  
CATCGAGGCCAGCAGCACCTGCTGCAAGCTGACCGTGTGGGCATCAAGCAGCTGCA  
GCGTGTGGCGTGGAGCGCTACCTGAAGGACCAAGCAGCTGCTGGGCATCTGGGCTG  
GGCAAGCTGATCTGCAACCACCGCGTCCCTGGAACGCCAGCTGGAGCAACAAGAGCTGG  
CCAGATCTGGAACAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACC  
TGATCTACACCCGATCGAGGGAGGCCAGAACACAGCAGGAGAACGAGCAGGAGCTG  
GAGCTGGACAAGTGGCCACCTGTTGAACTGTTGACATCAGCAAGTGGCTGTGGTACAT  
CAAGATCTCATCATGATCGTGGCGCCCTGGTGGGCCTGCGCATCGTGTACCGTGTGAG  
CATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCAAGCCGCTTCCCCGCC  
CCCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGAGCGCAGCGCAGCG  
AGCAGCCCCCTGGTGCAGGCCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTG  
TTCAGCTACCAACGCCCTGCGCAGCTGATCCTGATCGCCGCCGATCGTGGAGCTGCTGG  
CGCCGCCGCTGGAGGGCCCTGAAGTACTGGGCAACCTGCTGCACTGAGCTGGATCC  
GAAGAACAGGCCGTGAGCTGTCACGCCATGCCATGCCGTGGCGAGGGCACC  
GCATCATCGAGGTGGCCAGCGCATCGGCCGCCCTCCTGCACATCCCCGCCGATCCGCC  
AGGGCTCGAGCGCCCTGCTGTAACCGAG

FIG. 18

## SEQ ID NO:16 GLN422-TYR435

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGGCCGTGGAGAACGCTGGGTGACCGTGTACTACGGCGTCCCCGTG  
TGGAGGAGGCCACCAACCACCTGTTCTGCGCCAGCGACGCCAACGCCAACGGAGGT  
GCACAACCGTGTGGGCCACCCACGCCCTGCCGTGCCAACCGAACCCCAACCCCCAGGAGATCGTGCT  
GGAGAACGTGACCGAGAACCTCAACATGTGGAGAACACAACATGGTGGAGCAGATGCAAG  
GACATCATCAGCCTGTGGGACCAAGGCCATGAAAGCCCTGCGTGAGCTGACCCCCCTGCGTG  
ACCCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAGGAGAT  
GGACCGCGCGAGATCAAGAACCTGCAAGGTGACCCACCGCATCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTCCCCATGACAACGACAACACCAGC  
TACAAGCTGATCAACTGCAACACCCAGCGTGATCACCCAGGCCCTGCCAACGGTGGCTTCGA  
GCCCATCCCCATCCACTACTGCGCCCCGCCGGCTCGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGGGCCCTGACCAACGTGAGCACCCTGCACTGCACCCACGGCATCCGCC  
CCGTGGTGGAGCACCCAGCTGCTGCTGAACGGCAGCTGGCCAGGAGGGCGTGGTGAATCCGC  
AGCGAGAACCTCACCGACAAACGCCAACGACCATCATCGTCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCAACCGCCCCAACACAACACCCGCAAGAGCATCACCATCGGCCCCGGCGCCT  
TCTACGCCACCGGACATCATCGGCCACATCCGCCAGGCCACTGCAACATCAGCGGAG  
AAAGTGGAAACAACCCCTGAAGCAGATCGTACCAAGCTGCAAGGCCAGTTGGCAACAAAGAC  
CATCGTGTCAAGCAGAGCAGCGGGCGAACCCGAGATCGTGATGACAGCTTCAACTGCG  
GGGGCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAAACACCATCG  
CCCCCAACAAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGGGCGCTACGCC  
CCCCCCATCCCGGGCCAGATCCGCTGCGAGCAACATCACCGGCCCTGCTGCTGACCCCGAC  
GGCGGCAAGGAGATCAGCAACACCAACCGAGATCTCCGCCCCGGCGGGGACATGCGCGA  
CAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGGCCCTGGCGTGGCCC  
CCACCAAGGCCAACGCCAGCGCGCTGGTGAGCGCAGAACGCCAGCTGGCAGCCCTGGCGCATG  
TTCCTGGGCTTCTGGCGCCGCCAGCACCAGTGGCTGCTGAGCGCATCGTCAGCAGAACAAACCTGCTGCGGCCATCGA  
GGCCCAGCAGCACCTGCTGCACTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCGTGC  
TGGCGTGGAGCGCTACCTGAAGGACCAAGCAGCTGCTGGCATCTGGGCTGCAAGCGCAAG  
CTGATCTGCAACCACCGCGTGCCTGGAACGCCAGCTGGAGCAACAAAGAGCTGGACAGAT  
CTGGAACAAACATGACCTGGATGGAGTGGAGCGAGATCGACAACACTACACCAACCTGATCT  
ACACCCCTGATCGAGGGAGAGCCAGAACACAGCAGGAGAACGAGCAGGAGCTGCTGGAGCT  
GGACAAAGTGGCCAGCCTGTTGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGA  
TCTTCATCATGATCGTGGCGGCCCTGGCTGCTGGCATCGTCAGCTTCAACCGCTTCCCCCGCG  
TGAACCGCGTGCAGGCCAGGGCTACAGCCCCCTGAGCTTCAACCGCTTCCCCCGCG  
GCCCGACCGCCCCGAGGGCATCGAGGAGGGCGAGCGCAGCCGACCGCAGCG  
CCCCCTGGTGACGCCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGCGCTGTTCAAG  
CTACCAACCGCCTGCCGACCTGATCCTGATCGCCGCCGATCGTGGAGCTGCTGGCGCG  
CGGCTGGAGGCCCTGAAGTACTGGGCAACCTGCTGCACTGAGTCCAGGAGCTGAAGA  
ACAGCGCCGTGAGCCTGTTGACGCCATGCCATGCCGTGGCGAGGGCACCGACCGCATC  
ATCGAGGTGGCCCAGCGCATGGCCGCCCTGACATCCCCGCCGATCCGCCAGGGC  
TTCGAGCGGCCCTGCTGTAACCTGAG

FIG. 19

**SEQ ID NO:17 GLN422-TYR435B**

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGCGCCGTGGAGAAGCTGTTGGTGACCGTGTACTACGGCGTCCCCGTG  
TGGAGGAGGCCACCACCACTGCTGCCAGCGACGCCAAGGCCTACGGACACCCGAGGT  
GCACAACGTGTGGCCACCCACGCCCTGCGCCACCGACCCCAACCCCAAGGAGATCGTGT  
GGAGAACGTGACCGAGAACCTCAACATGTGGAGAACACACATGGTGGAGCAGATGCACGAG  
GACATCATCGCCTGTGGGACCAAGAGCTGAAGCCCTGCGTGAAGCTGACCCCCCTGCGTGT  
ACCCCTGCACCTGCAACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAGGAGAT  
GGACCGCGCGAGATCAAGAACCTGCAAGGTGACCCACCGCATCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTTCAAAAGCTGGACGTGGTCCCCATCGACAACGACAACACCCAGC  
TACAAGCTGATCAACTGCAACACCCAGCGTGTACCCAGGCCCTGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCCTGAGTGCACCCACGGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGTAACGGCAGCCCTGCCAGGAGGGCGTGGTATCCGC  
AGCGAGAACCTCACCGACAACGCCAACGACCATATCGTCAAGGAGAGCGTGGAGAT  
CAACTGCACCGCCCCAACAAACAACACCCGAAGAGCATACCATCGGCCCCGGCCGCGCCT  
TCTACGCCACCGCGACATCATCGGCACATCCGCCAGGCCACTGCAACATCAGCGCAG  
AAAGTGGAACAAACACCCAGCTGAGCAGATCGTGAACAGCTGAGGCCAGTTGGCAACAAAGAC  
CATCGTGTCAAGCAGAGCAGCGCCGACCCAGATCGTGTATGCACAGCTCAACTGCG  
GCCGCGAGTTCTACTGCAACAGCACCCAGCTGTTAACAGCACCTGGAACAAACACCATCG  
GCCCAACAAACACCAACGGCACCATCACCTGCCCTGCCATCAAGCAGGCCCTACGCC  
CCCCCATCCGCGCCAGATCCGCTGCAAGCAGCAACATCACCGCCCTGCTGCTGACCCGCGACG  
GCCGCAAGGAGAGATCAGCAACACCCAGAGATCTTCCGCCCCGGCGCGACATGCGCAG  
AACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGGCCCTGGCGTGGCCC  
CACCAAGGCCAAGGCCCGCGTGGTGCAGCGAGAACGCCAGCTGACCTGGCGCCATGT  
TCCTGGGCTTCCCTGGCGCCCGCGCAGCACCATGGCGCCCGCAGCCTGACCTGACCGTGC  
AGGCCGCGCAGCTGCTGAGCGGCATCGTCAAGCAGAGAACACCTGCTGCGCGCCATCGAG  
GCCCAAGCAGCACCTGCTGCACTGACCGTGTGGGCATCAAGCAGCTGCAAGGCCCGCTGCT  
GCCGTGGAGCGTACCTGAAGGACAGCAGCTGCTGGCATCTGGGCTGAGCGGCAAGC  
TGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCTGGACCAGATC  
TGGAAACAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAACACTACCAACCTGATCTA  
CACCTGATCGAGGAGAGCCAGAACCCAGCAGGAGAACAGAGCAGGAGCTGCTGGAGCTG  
GACAAGTGGGCCAGCCTGTGGAACCTGGTGCACATCAGCAAGTGGCTGTGGTACATCAAGAT  
CTTCATCATGATCGTGGCGCCAGGGCTACAGCCCCCTGAGCTTCAAGCAGCTGCTTACCGTGTGAGCATCGT  
GAACCGCGTGCAGGCCAGGGCTACAGCCCCCTGAGCTTCAAGCAGCTGCTTACCGTGTGAGCATCGT  
CCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGAGCGCGACCGCGACCGCAGCAGC  
CCCCTGGTGCAGGCCCTGCTGGCCCTGATCTGGGAGCAGCTGCGCAGCCTGCTGCTTACCGT  
TACCAACGCCCTGCGCGACCTGATCCTGATCGCCGCCGCACTCGTGGAGCTGCTGGGCCCGC  
GGCTGGAGGCCCTGAAGTACTGGGCAACCTGCTGCACTGAGTACTGGATCCAGGAGCTGAAGAA  
CAGCGCCGTGAGCTGTTGACGCCATGCCATGCCATGCCGCCCTCCTGCACATCCCCGCC  
CGAGCGGCCAGCCATGCCGCCATGCCGCCATGCCGCCAGGGCT  
CGAGCGGCCCTGCTGTAACCTGAG

**FIG. 20**

## SEQ ID NO:18: LEU122-SER199; ARG426-GLY431

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTCGCCCAGCGCCGTGGAGAACGCTGTGGGTGACCGTGTACTACGGCGTGCCGTG  
TGGAGGAGGCCACCAACCCCTTCTGCAGCCAGCGAACGCCAACGCCAACGGAGGT  
GCACAAACGTGTGGGCCACCCACGCGTGCCTGCCAACGCCAACCCCCAGGAGATCGTGC  
GGAGAACGTGACCGAGAACCTCAACATGTGGAGAACAAACATGGTGGAGCACAGTCACGAG  
GACATCATCAGCTGTGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGCAACAGCGTGT  
CACCCAGGCCCTGCCAACGGTGAAGCTTCAGCCCATCCACTACTGCGCCCCCGCCGG  
CTTCGCCATCCTGAAGTGAACGACAAGAACGTTCAACGGCAGCGCCCCCTGCACCAACGTGA  
GCACCGTGCAGTGCACCCACGGCATCCGCCCGTGGTGAAGCAGCCAGCTGCTGTAACGGC  
AGCCTGGCCGAGGAGGGCGTGGTGAATCCGAGCGAGAACCTCACCGACAACGCCAACGAC  
CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAAACACCCGCA  
AGAGCATACCATCGGCCCGGCCGCGCCTTACGCCACCGGCACATCATCGCGACATCC  
GCCAGGCCACTGCAACATCAGCGGAGAACAGTGGAACAAACACCCCTGAAGCAGATCGTGC  
AAGCTGCAGGCCAGTTCGGAACAAAGACCATCGTGTCAAGCAGAGCAGGGCGGCGACCC  
CGAGATCGTGTGACAGCTCAACTCGCGGGCAGTTCTACTGCAACAGCACCCAGCT  
GTTCAACAGCACCTGAAACAACACCATCGGCCAACAAACACCAACGGCACCATCACCC  
CCTGCCGCATCAAGCAGATCATCAACCGCGGCCGCAAGGCCATGTACGCCCGCCATCC  
GCCGCCAGATCCGCTGCAGCAGAACATCACCGGCCGCTGCTGACCCCGCACGGCGCAAG  
GAGATCAGCAACACCAACCGAGAGATCTCCGCCCGGCCGACATGCGCACAACTGGCG  
CAGCGAGCTGTACAAGTACAAGGTGGTAAGATCGAGCCCTGGCGTGGCCCCAACCAAGG  
CCAAGCGCCGCGTGGTGCAGCGAGAACAGCGCCCGTGAACCTGGCGCCATGTCCTGGC  
TTCCTGGCGCCGCCGAGCACCATGGCGCCCGCAGCCTGACCCCTGACCGTGCAGGCCGC  
CAGCTGCTGAGCGGCATCGTGCAGCAGCAAAACACTGCTGCGGCCATCGAGGCCAGCA  
GCACCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGTGTGGCGTGG  
AGCGTACCTGAAGGACCAACAGCAGCTGCTGGGCATCTGGCTGAGCGCAAGCTGATCTGC  
ACCACGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGAAACAA  
CATGACCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACCAACCTGATCTACACCC  
TCGAGGAGAGCCAGAACACAGCAGGAGAACAGCAGCAGGAGCTGTTGGAGCTGGACAAGTG  
GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT  
GATCGTGGCGGCCCTGGTGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTAACCGCGT  
GCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCGGCCCCGACCG  
CCCCGAGGGCATCGAGGAGGGAGGGCGAGCGCAGCCGACCGCAGCAGCCCCCTGGTGC  
ACGGCCTGCTGGCCCTGATCTGGAGCAGACCTGCGCAGCCTGTGCCTGTCAGCTACCC  
TGCAGCAGCTGATCTGATGCCGCCGATCGTGGAGCTGCTGGCTGGCCGCCGGCTGGAGG  
CCCTGAAGTACTGGGCAACTGCTGCACTGAGCTACTGGATCCAGGAGCTGAAGAACAGGCC  
AGCCTGTCGACGCCATGCCATGCCGTGGCCAGGGCACCGACCGCATCATCGAGGTGGCC  
CAGCGCATCGGCCGCCCTTCTGCACATCCCCGCCATCCGCCAGGGCTCGAGCGCGCC  
CTGCTGTAACCGAG

FIG. 21

SEQ ID NO:19 LEU122-SER199; ARG426-LYS432

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTCGTTGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAGGAGGCCACCAACCACCTGTCGCCAGCGACGCCAAGGCCACGACACCGAGGT  
GCACAAACGTGTGGGCCACCCACGCCCTGCGTGCCTGCCACCGACCCCAACCCCCAGGAGATCGTGCT  
GGAGAACGTGACCGAGAACTTCAACATGTGGAGAAACAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAGGCCCTGAAGGCCCTGCGTGAAGCTGGCAACAGCGTGAT  
CACCCAGGCCCTGCCCAAGGTGAGCTCGAGGCCATCCCCATCCACTACTGCGCCCCCGCG  
CTTCGCCATCCTGAAGTGCACGACAAGAACGTTCAACGCCAGCGGGCCCTGCACCAACGTGA  
GCACCGTGCAGTGACCCACGGCATCCGCCCGTGGTGAAGCAGCTGCTGCTGAACGGC  
AGCCTGGCCGAGGAGGGCGTGGTATCCGCAGCGAGAACCTCACCGACAACGCCAAGACCAT  
CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAACACCCGCA  
AGAGCATACCATCGGCCCGGCCCTACGCCACCGCGACATCATCGCGACATCC  
GCCAGGCCACTGCAACATCAGCGCGAGAACGAGTGGAAACAACACCCGTAAGCAGATCGGAC  
AAGCTGCAGGCCAGTTCGGCAACAAGACCATCGTGTCAAGCAGAGCAGGGCGGCCACCC  
CGAGATCGTGTATGCACAGCTCACTCGGGCGGGAGTTCTACTGCAACAGACCCAGCT  
GTTCAACACGACCTGAAACAACACCATCGGCCAACAACACCAACGGCACCATCACCGT  
CCTGCCGCATCAAGCAGATCATCAACCGCGCGCAACAAGGCCATGTAACGCCCGCCATCC  
GCCAGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCCTGCTGCTGACCCGCGACGGCG  
GAGATCAGCAACACCAACCGAGATCTCCGCCCGCGGCCGACATGCGCGACAACCTGGC  
CAGCGAGCTGTACAAGTACAAGGTGGTAAGATCGAGCCCTGGCGTGGCCCCCACCAGG  
CCAAGCGCCGCGTGGTGCAGCGCGAGAACGCGCCCGTGAACCTGGCGCAATGTTCTGGC  
TTCCTGGCGCCGCCGAGCACCATGGCGCCCGCAGCCTGACCCCTGACCGTGCAGGGCCGC  
CAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACACCTGCTGCCGCCATCGAGGCCAGCA  
GCACCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGTGTGGCGTGG  
AGCGCTACCTGAAGGACCAGCAGCTGCTGGCATCTGGGCTGCAGCGCAAGCTGATCTGC  
ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAAGATCTGGAAACAA  
CATGACCTGGATGGAGTGGAGCGCGAGATCGACAACACTACACCAACCTGATCTACACCGT  
TCGAGGAGAGCCAGAACACAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG  
GGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT  
GATCGTGGCGGCCCTGGTGGCCTGCGCATCGTGTTCACCGTGTGAGCATCGTAACCGCGT  
GCGCCAGGGCTACAGCCCTGAGCTTCAGACCCGTTCCCGCCCGCGGCCGACCG  
CCCCGAGGGCATCGAGGAGGGCGCGAGCGCAGCCGACCGCAGCACGCCCTGGTGC  
ACGGCCTGCTGGCCCTGATCTGGACGACCTGCCAGCCTGCTGCTGTCAGTACCAACCGCC  
TGCAGCAGCTGATCTGATGCCGCCGATCGTGGAGCTGCTGGGCCCGCGCTGGAGG  
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG  
AGCCTGTTGACGCCATGCCATGCCGTGGCCAGGGCACCGACCCGATCATCGAGGTGGCC  
CAGCGCATCGGCCGCCCTCCTGCACATCCCCGCCGATCCGCCAGGGCTCGAGCGCGCC  
CTGCTGTAACCGAG

FIG. 22

SEQ ID NO: 20: LEU122-SER199; TRP427-GLY431

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGTGGAGCA  
GTCTTCGTTGCCAGCGCCGTGGAGAAGCTGTTGACCGTGTACTACGGCGTCCCCGTG  
TGGAGGAGGCCACCAACCCTGTTCTGCGCCAGCGACGCCAAGGCCAACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCCTGCGTGCACCGACCCCAACCCCCAGGAGATCGTGT  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCTGAAGGCCCTGCGTGAAGCTGGCAACAGCGTGT  
CACCCAGGCCATGCCAAGGTGAGCTCGAGCCCATCCCCATCCACTACTGCGCCCCCGGG  
CTTCGCCATCCGTGAAAGTCAACGACAAGAACGTTCAACGGCAGCGGGCCCTGCAACAAACGTGA  
GCACCGTGCAGTGCACCCACGGCATCCGCCCGTGGTGAATCCGCAGCGAGAACCTACCGACAACGCCAAGACCAT  
CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAAACACCCGCA  
AGAGCATCACCATCGGCCCGGCCCTTACGCCAACGGCAGATCGCGCACATCC  
GCCAGGCCACTGCAACATCAGCGCGAGAACGAAACACCCCTGAAGCAGATCGTGT  
AAGCTGCAAGGCCAGTCGGCAACAAGACCATCGTGTCAAGCAGAGCAGCGGGCGACCC  
CGAGATCGTGTGACAGCTCAACTCGGGCGCGAGTTCTACTGCAACAGCACCCAGCT  
GTTCAACAGCACCTGGAACAAACACCATCGGCCAACAAACACCAACGGCACCACCATC  
CCTGCCGCATCAAGCAGATCATCAACCGCTGGGCGGCAAGGCCATGTACGCC  
GCCGCCAGATCCGCTGCAGCAGCAACATCACCGCCCTGCTGTCACCCCGCAGCGGGCAAG  
GAGATCGCAACACCACCGAGATCTTCCGCCCGCGCGGACATCGCGACA  
CAGCGAGCTGTACAAGTACAAGGTGGTAAGATCGAGCCCTGGCGTGGCCCCAACCAAGG  
CCAAGCGCCGCGTGGTCAGCGAGAACGCGCCGTGACCCCTGGCGCCATGTTCTGGC  
TTCCCTGGCGCCGCCGGCAGCACCATGGCGCCCGCAGCCTGACCCCTGACCGTGCAGGCC  
CAGCTGCTGAGCGGCATCGTCAGCAGCAGAACAAACCTGCTGCGGCCATCGAGGCC  
GCACCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCGAGCCCGTGTGGCG  
AGCGCTACCTGAAGGACCAGCAGCTGCTGGCATCTGGGCTGCGAGCGAACGCTG  
ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGAACAAAGAGCCTGGACCAGATCTGGAA  
CATGACCTGGATGGAGTGGGAGCGCAGATCGACA  
ACTACACCAACCTGATCTACACCC  
TCGAGGAGAGCCAGAAC  
CAGCAGGAGAGAACGAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG  
GGCCAGCCTGTGGA  
ACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTCAT  
GATCGTGGCGGCC  
CTGGTGGCCTGCGCATCGTGT  
GCGCCAGGGCTACAGCCCC  
CTGAGCTCCAGACCCGCTTCCCC  
GCCGGCGGCC  
CCCCGAGGGCATCGAGGAGGGCGGAGCGCGACCC  
CGCAGCAGGCC  
ACGGCCTGCTGGCC  
CTGATCTGGAGCAGACCTGCGCAGCCTG  
TGC  
CGCAGCTGAT  
CTGACCGGCC  
CATCGTGGAGCTGCTGGCG  
CCCTGAAGTACTGGGCAAC  
CTGCTGCA  
GAGTACTGGATCCAGGAGCTGAAGAACAGCGCC  
AGCCTGTTGACGCC  
ATCGCCATCGCGTGGCGAGGG  
CACCGACCG  
CATCGAGGTGGCC  
CAGCGC  
ATCGGCCGCC  
CTTCC  
TGACATCCGCC  
AGGGCTTC  
GAGCGCGCC  
CTGCTGTA  
ACTCGAG

FIG. 23

SEQ ID NO:21 LYS121-VAL200; ASN425-LYS432

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCCAGCGCCGTGGAGAACGCTGTGGGTGACCGTGTACTACGGCGTGCCTCGT  
TGGAGGAGGCCACCAACCCCTGTCGCCAGCAGCCAAGGCCAACGGACACCCGAGGT  
GCACAACGTGTGGCCACCCACGCCCTCGTGCCTGCCACCGACCCCCAACCCCCCAGGAGATCGTGC  
GGAGAACGTGACCGAGAACCTCAACATGTGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCCTGTCGCCACCGACCCCCAACCGACACCCGAGGT  
GCCCTGCCCAAGGTGAGCTCGAGGCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGC  
CATCCTGAAGTGAACGACAAGAACGTTCAACGGCAGCGGGCCCTGCAACCGACACCCGAGGT  
TGCAGTGCACCCACGGCATCCGCCCGTGGTGAACGCCAGCTGCTGTCACGGCAGCCTGG  
CCGAGGAGGGCGTGGTGAATCCGCAAGCAGAACCTCACCGACAACGCCAACGACCATATCGT  
CAGCTGAAGGAGACGTGGAGATCAACTGCACCCGCCAACAACACACCCGCAAGAGCAT  
CACCATCGGCCCCGGCGGCCCTCTACGCCACCGCGACATCATCGCGACATCCGCCAGGC  
CCACTGCAACATCAGCGCGAGAACGTTCAACACCCCTGAAGCAGATCGTGAACGCCAGCTGC  
AGGCCAGTTCGGCAACAAGACCATCGTGTTCAGCAGAGCAGCAGCGGGCAACCCGAGATC  
GTGATGCAAGCTCAACTGCCGCCGAGTTCTCTACTGCAACAGCACCCAGCTGTTCAAC  
AGCACCTGGAACAAACACCATCGGCCAACAACACCAACGGCACCATCACCCCTGCCCTGCC  
CATCAAGCAGATCATCAACGCCCAAGGCCATGTACGCCCTGCCATCCGCCAGATCCG  
CTGCAGCAGCAACATCACCGCCCTGCTGTCACCGCGACGGCGCAAGGAGATCAGCAACA  
CCACCGAGATCTTCGCCCGGGCGGCCGACATCGCGACAACTGGCGCAGCGAGCTGTAC  
AAAGTACAAGGTGGTGAAGATCGAGCCCTGGCGTGGCCCCAACCAAGGCCAGCGCCG  
GGTGCAGCGCGAGAACGCGCCGTGACCCCTGGCGCAATGTTCTGGCTTCTGGCAGCTGCC  
CGGCAGCACCATGGCGCCCGAGCCTGACCCCTGACCGTGCAGGCCAGCTGCTGAGCG  
GCATCGTCAGCAGCAGAACACCTGCTGCCGCCATCGAGGCCAGCAGCACCTGCTGCAG  
CTGACCGTGTGGGCATCAAGCAGCTGCAAGGCCCGTGTGGCCGGTGGAGCGCTACCTGAA  
GGACCAAGCAGCTGCTGGCATCTGGGCTGCAAGCAGAACGCTGATCTGCACCAACCGCG  
CCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAAACATGACCTGGATG  
GAGTGGGAGCGCGAGATCGACAACACCAACCTGATCTACACCTGATCGAGGAGAGCCA  
GAACCAAGCAGGAGAAAGAACGAGCAGGAGCTGCTGGAGCTGACAAGTGGCCAGCCTGTGG  
AACTGGTTGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGCG  
CTGGTGGGCCTGCGCATCGTGTTCACCGTGTGAGCATCGTAACCGCGTGCAGGCCAGGGCTAC  
AGCCCCCTGAGCTTCAGACCCGCTCCCCGCCCGGGCCCCGACCGCCCCGAGGGCATC  
GAGGAGGAGGGCGCGAGCGCGACCGCGACCGCAGCACGCCCTGGTGCACGCCCTGCTGGC  
CCTGATCTGGACGACCTGCGCAGCCTGTCCTGTTCACTGACCCCGCTGCCGACCTGAT  
CCTGATGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGGCTGGAGGCCCTGAAGTACTG  
GGCAACCTGCTGCACTGGATCCAGGAGCTGAAAGAACAGCGCCGTGAGCCTGTCAG  
CCATGCCATGCCGTGGCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGC  
CGCGCCTTCTGCAACATCCCCGCCGATCCGCCAGGGCTCGAGCGCGCCCTGCTGTAAC  
GAG

FIG. 24

## SEQ ID NO:22 VAL120-ILE201; ILE 424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTCGTTGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTCCCCTG  
TGGAGGAGGCCACCAACCACCCCTGTTCTGCGCCAGCGACGCCAAGGCCCTACGACACCGAGGT  
GCACAAACGTGTGGCCACCCACGCCCTGCGTCCCACCGACCCCAACCCCAAGGAGATCGTGT  
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGACCTGAAGCCCTGCGTGGGCGGATCACCCAGGCCTG  
CCCCAAGGTGAGCTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTCGCCATCCT  
GAAGTGAACGACAAGAACGTTCAACGGCAGCGGCCCTGACCAACGTGAGCACCCTGCAGT  
GCACCCACGGCATCCGCCCGTGGTAGCACCCAGCTGCTGTAACGGCAGCCTGGCGAG  
GAGGGCGTGGTATCCGAGCGAGAACGTTCAACCGACAACGCCAAGACCATATCGTGCAGCT  
GAAGGAGAGCGTGGAGATCAACTGCAACCGCCCCAACAAACAACACCCGCAAGAGCATCACCA  
TCGGCCCCGGCCGCCCTACGCCACCGGACATCATCGCGACATCCGCCAGGGCCACT  
GCAACATCAGCGGCAGAACGTTCAACACCCCTGAAGCAGATCGTGAACAGCTGCAGGCC  
CAGTTCGCAACAAGACCATCGTGTCAAGCAGAGCAGCGGCCAGCCGAGATCGTGT  
GCACAGCTCACTGCGGCCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC  
CTGGAACAACACCATCGGCCCAACAAACACCAACGGCACCACCATACCCCTGCCGCATCA  
AGCAGATCATCGGGCGCCATGTACCCCCCCCCATCCGCGCCAGATCCGCTGCAGCAGC  
AACATCACCGGCTGCTGCTGACCCGGACGGCGCAAGGAGATCAGCAACACCAACCGAGAT  
CTTCGCCCGGGCGGCCACATGCGGACAACTGGCGAGCAGCTGTACAAGTACAAGG  
TGGTGAAGATCGAGCCCTGGCGTGGCCCCACCAAGCCAAGCGCCGCGTGGTGCAGCGC  
GAGAACGCGCGCCGTGACCCCTGGCGCCATGTTCTGGCTTCTGGCGCCGCCAGCACC  
ATGGCGCCCGCAGCCTGACCCCTGACCGTGACGGCCGCCAGCTGAGCGGCATCGTCA  
GCAGCAGAACAACTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCGTGT  
GGGGCATCAAGCAGCTGAGGCCGCGTGGAGCGCTACCTGAAGGACCAAGCAG  
CTGCTGGCATCTGGGCTGAGCGCAAGCTGATCTGACCAACCCGTGCCCTGGAACGCC  
AGCTGGAGCAACAAAGACCTGGACCAAGATCTGGAACAAACATGACCTGGATGGAGTGGAGCG  
CGAGATCGACAACACCAACCTGATCTACACCCCTGATCGAGGAGAGCCAGAACACAGCAG  
AGAAGAACGAGCAGGAGCTGGAGCTGGACAAGTGGCCAGCCTGGAACTGGTCAC  
ATCAGCAAGTGGCTGTTACATCAAGATCTTACATGATCGTGGCGCCCTGGTGGCCCTG  
CGCATCGTGTACCGTGTGAGCATGTCGAACCGCGTGCACGCCAGGGCTACAGCCCCCTGAGC  
TTCCAGACCCGCTCCCCGCCCGGCCAGGCCAGGGCATCGAGGAGGG  
CGCGAGCGCGACCGCGACCGCAGCACGCCCTGGTGCACGCCCTGCTGGCCCTGATCTGG  
ACGACCTGCGCAGCCTGTCCTGAGCTACCAACCGCCCTGCGCAGCCTGATCCTGATCGCCG  
CCCGCATCGTGGAGCTGCTGGGCCCGCGCTGGAGGCCCTGAAGTACTGGGGCACCTG  
CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATGCCATC  
GCCGTGGCGAGGGCACCGACCGCATCGAGGTGGCCAGCGCATGGCCGCGCCTTC  
GCACATCCCCGCCGATCCGCCAGGGCTCGAGCGCGCCCTGCTGTAACCGAG

FIG. 25

SEQ ID NO:23: VAL120-ILE201B; ILE424-ALA433

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTGCCAGCGCCGTGGAGAACGCTGTGGGTGACCGTGACTACGGCGTGCCTG  
TGGAGGAGGCCACCACCACTGTTCTGCAGCCAAGGCCAACGACACCGAGGT  
GCACAACCGTGTGGGCCACCCACGCCTGCCTGCCACCGACCCCCAACCCCCAGGAGATCGTG  
GGAGAACGTGACCGAGAACCTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCTGTGGGACCAGAGCCTGAAGCCCTGCCTGCCGGCATCACCCAGGCGTGC  
CCCAAGGTGAGCTCGAGCCCCATCCCCATCCACTACTGCCTGCCGGCTCGCATCCTG  
AAAGTCAACGACAAGAACGTTCAACGGCAGCGCCCCCTGCACCAACGTGAGCACCGTGCAGTG  
CACCCACGGCATCCGCCCCGTGGTGACCAACCCAGCTGCTGTAACGGCAGCCTGGCCGAGG  
AGGGCGTGGTGTCCGAGAACCTCACCGACAACGCCAACGACATCGTGAGCTG  
AAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAACAAACACCCGCAAGAGCATCACCAT  
CGGCCCGGCCGCGCTTCTACGCCACGGCAGCATCGCGACATCCGCCAGGCCACTG  
CAACATCAGCGCGAGAACGTTCAACGAGCAGCAGCGCCGACCCCCGAGATCGTGATG  
AGTTCGGCAACAAGACCATCGTGTCAAGCAGCAGCGCCGACCCCCGAGATCGTGATG  
CACAGCTCAACTGCAGCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACC  
TGGAAACAACACCATCGCCCCAACAACACCAACGGCACCACCATCACCTGCCCTGCCGCATCAA  
GCAGATCATCGCGGCCATGTACGCCCGCCATCGCGCCAGATCCGCTGCAGCAGCA  
ACATCACCGGCCTGCTGCTGACCCCGCAGCGCCAGGAGATCAGCAACACCCACCGAGATC  
TTCCGCCCGCGCGCGACATCGCGACACTGGCGCAGCGAGCTGTACAAGTACAAGGT  
GGTGAAGATCGAGCCCCCTGGCGTGGCCCCCACCAAGGCCAACGCCGCGTGGTGCAGCGCG  
AGAACGCGCCGTGACCCCTGGCGCCATGTTCTGGCTTCTGGCGCCGCCAGCAGCACCA  
TGGCGCCCGCAGCCTGACCCCTGACCGTGAGGCCAGCTGCTGAGCGGCATCGCAG  
CAGCAGAACAACTGCTGCAGGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCCTGTG  
GGGCATCAAGCAGCTGAGGCCCGCTGCTGGCGTGGAGCGCTACCTGAAGGACCAGCAGC  
TGCTGGCATCTGGGCTGAGCGCAAGCTGATCTGCACCAACGCCGTGCCCTGGAACGCCA  
GCTGGAGCAACAAGAGCCTGGACAGATCTGAAACAAACATGACCTGGATGGAGTGGAGCGC  
GAGATCGACAACACTACACCAACCTGATCTACACCCCTGATCGAGGAGAGCCAGAACCCAGCAGGA  
GAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGCCAGCCTGGAACCTGGTCA  
TCAGCAAGTGGCTGTTACATCAAGATCTTCATCATGATCGTGGCGGCCCTGGTGGCCTG  
GCATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGCCTGCCAGGGCTACAGCCCCCTGAGCT  
TCCAGACCCGCTTCCCCGCCCGGCCGACCGCCAGCAGGCCCTGGTGCACGGCTGCTGGCCCTGATCTGGGA  
GGCGAGCGCAGCCGACCGCAGCAGGCCCTGGTGCACGGCTGCTGGCCCTGATCTGGGA  
CGACCTGCGCAGCCTGTCCTGTTAGCTACCAACGCCCTGCCGCGACCTGATCTGATGCCGC  
CCGCATCGTGGAGCTGCTGGCCGCCGGCTGGAGGCCCTGAAGTACTGGGCAACCTGC  
TGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATGCCATC  
GCCGTGGCCGAGGGCACCGACCGCATCTGAGGTGGCCAGCGCATCGGCCGCGCTTCC  
GCACATCCCCGCCATCCGCCAGGGCTCGAGCGCCCTGCTGTAACCTGAG

FIG. 26

**SEQ ID NO:24 VAL120-THR202; ILE424-ALA433**

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
 GTCTTCGTTGCCCGCGCCGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
 TGGAGGAGGCCACCACCACTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT  
 GCACAAACGTGTGGGCCACCCACGCCCTGCGTGCCTGCCACCGACCCCAACCCCAAGGAGATCGTGT  
 GGAGAACGTGACCGAGAACCTCAACATGTGGAGAACAAACATGGTGGAGCAGATGCACGAG  
 GACATCATCAGCCTGTGGGACCCAGGCCCTGAAGGCCCTGCGTGGGCCACCCAGGCC  
 CCCCAAGGTGAGCTCGAGCCATCCCCATCCACTACTGCGCCCCCGCCGGCTCGCCATCCT  
 GAAGTGCAACGACAAGAACCTCAACGGCAGCGCCCTGCACCAACGTGAGCACCCTGCA  
 GCACCCACGGCATCGCCCCGTGGTGGAGCACCCAGCTGCTGTAACGGCAGCCTGGCCAG  
 GAGGGCGTGGTGAATCCGAGCGAGAACCTCACCGACAACGCCAAGACCATATCGTGAGCT  
 GAAGGAGAGCGTGGAGATCAACTGCACCCGCCAACAACAAACACCCGCAAGAGCATCACCA  
 TCGGCCCCGGCGCGCTTCTACGCCACCGCGACATCATCGCGACATCCGCCAGGCCACT  
 GCAACATCAGCGGCAGAACAGTGGAACAAACACCCCTGAAGCAGATCGTACCAAGCTGCAGGCC  
 CAGTCGGCAACAAGACCATCGTGTCAAGCAGAGCAGCGCCGGCACCCAGAGATCGTGAT  
 GCACAGCTTCAACTCGGGCGCGAGTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC  
 CTGGAACAAACACCATCGGCCAACAACACCAACGGCACCATCACCTGCCCGCATCA  
 AGCAGATCATCGGCGGCATGTACGCCCGGCGACATCGCGGAGCTGAGCAGCAGC  
 AACATCACCGGCCTGCTGTCACCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGAT  
 CTTCCGCCCGGCGCGGCGACATCGCGGAGCTGAGCAGCTGAGCTGAGCAG  
 TGGTGAAGATCGAGGCCCTGGCGTGGCCCGAACCAAGGCCAAGCGCCGCGTGGTCA  
 GAGAAGCGCGCCGTGACCTGGCGCCATGTTCTGGCTTCTGGCGCCGCCGGCAGCAC  
 ATGGGCGCCCGCAGCCTGACCTGACCGTGAGGCCCGCCAGCTGCTGAGCGGCA  
 GCAGCAGAACAAACCTGCTGCCGCAATCGAGGCCAGCAGCACCTGCTGAGCTGACCGTGT  
 GGGGCATCAAGCAGCTGCAAGGCCCGCGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG  
 CTGCTGGGATCTGGGCTGCAAGCGCAAGCTGATCTGCACCAACCGCGTGCCCTGGAACGCC  
 AGCTGGAGCAACAAGAGCCTGGACCGAGATCTGGAACAAACATGACCTGGATGGAGTGGAGCG  
 CGAGATCGACAACACCAACCTGATCTACACCCCTGATCGAGGAGAGCCAGAACCGAGCAG  
 AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGCCAGCCTGGAACCTGGTCA  
 ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGCGGCCCTGGTGG  
 CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCCTGCCAGGGCTACAG  
 TTCCAGACCGCTTCCCCGCCCGCGGCCCCGACCGCCCGAGGGCATCGAGGAGGG  
 CGCGAGCGCAGCCCGACCGCAGCAGGCCCTGGTGCACGCCCTGCTGCCCTGATCTGG  
 ACGACCTGCGCAGCCTGTGCCCTGTCAGCTACCAACCGCTGCGCAGCTGATCTG  
 CCCGCATCGTGGAGCTGCTGGCCGCCCGCTGGAGGCCCTGAAGTACTGGGCAACCTG  
 CTGAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTCACGCCATGCCATC  
 GCCGTGGCCGAGGGCACCGACCGCATATCGAGGTGGCCCAAGCGCATTGCCCGCCT  
 GCACATCCCCGCCGATCCGCCAGGGCTCGAGCGGCCCTGCTGTAACCTCGAG

**FIG. 27**

## SEQ ID NO:25 VAL127-ASN195

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCTGTG  
TGGAGGAGGCCACCAACCCCTGTTCTGCGCCAGCGACGCCAAGGCCACGACACCCGAGGT  
GCACAAACGTGTGGCCACCCACGCCGTGCCACCGACCCCCAACCCCCAGGAGATCGTGT  
GGAGAACGTGACCGAGAACTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGACCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGT  
GGGCAGGGAACTGCAACACCAGCGTGAACCCAGGCCGCCCCAAGGTGAGCTCGAGCC  
CATCCCCATCCACTACTGCGCCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT  
CAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGAGTCACCCACGGCATCCGCCCG  
TGGTGGAGCACCCAGCTGCTGTAACGGCAGCCTGGCCAGGAGGGCGTGGTATCCGAGC  
GAGAACATTCAACGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA  
CTGCACCCGCCAACAACACACCCGAAGAGCATCACCATCGGCCCCGGCGCCCTCTA  
CGCCACCGCGACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGCGAGAAC  
GGAACAAACACCCCTGAAGCAGATCGTACCAAGCTGCAAGGCCAGTTGGCAACAAAGACATC  
GTGTTCAACGAGCAGCGGGCGACCCGAGATCGTGTGACAGCTTCAACTGCCGG  
CGAGTTCTCTACTGCAACAGCACCCAGCTGTCACAGCACCTGGAACAAACACCATCGGCC  
CAACAAACACCAACGGCACCATCACCTGCCCTGCCATCAAGCAGATCATCAACCGCTGCC  
AGGAGGTGGCAAGGCCATGTACCCCCCCCCATCCGCGCCAGATCCGCTGCAAGCAGAAC  
ATCACCGGCCTGCTGTCACCGCGACGGCGCAAGGAGATCAGCAACACCAGAGATCTT  
CCGCCCCGGCGCGACATCGCGACAACTGGCGCAGCAGCTGACAAGTACAAGGTGG  
TGAAGATCGAGCCCCCTGGCGTGGCCCCCACAAGGCCAAGCGCCGCTGGTGCAGCGAG  
AAGCGGCCGTGACCTGGCGCCATGTTCTGGCTTCTGGCGCCGCCAGCACCAG  
GGCGCCCGCAGCCTGACCCGACCGTGCAAGGCCAGCTGCTGAGCGGCATCGTCA  
GCAGAACAAACCTGTCGCGCCATCGAGGCCAGCACCTGTCAGCTGACCGTGTGG  
GCATCAAGCAGCTGCAAGGCCCGTGTGGCCCGTGGAGCGTACCTGAAGGACCAAGCAGCTG  
CTGGCGATCTGGGGCTGCAAGCTGATCTGACCAACCGCCGTGCCCTGGAACGCCAG  
CTGGAGCAACAAAGGCCCTGGACAGATCTGGAACAAACATGACCTGGATGGAGTGGAGCG  
AGATCGACAACCTACACCAACCTGATCTACACCCCTGATCGAGGAGAGCCAGAACCGAG  
AAGAACGAGCAGGAGCTGGAGCTGGACAAGTGGCCAGCCTGGAACGGTTGACAT  
CAGCAAGTGGCTGTTACATCAAGATCTTCACTGATCGTGGCGCCCTGGTGGCG  
CATCGTGTTCACCGTGTGAGCATCGTAACCGCGTGCGCCAGGGCTACGCCCTGAGCTT  
CCAGACCCGCTCCCCGCCCGGGCCCCGAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGAC  
GGAGCGCGACCGCGACCGCAGCACGCCCCCTGGTGCACGGCCTGCGCACCTGATCTGATGCCGCC  
GACCTGCGCAGCCTGTGCCTGTTCACTGACCCGCGTGCAGCTGACCTGATCTGATGCCGCC  
CGCATCGTGGAGCTGCTGGCCGCCGGCTGGAGGCCCTGAAAGTACTGGGCAACCTGCT  
GCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATGCCATCG  
CCGTGGCCGAGGGACCGACCGCATCGAGGTGGCCAGCGCATCGGCCGCCCTCTGC  
ACATCCCCGCCATCCGCCAGGGCTCGAGCGCCCTGCTGTAACCGAG

FIG. 28

SEQ ID NO:26 VAL127-ASN195; ARG426-GLY431

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGGAGCA  
GTCTTCGTTTCGCCAGGCCGTGGAGAAGCTGTTGACCGTGTACTACGGCGTGCCCCTG  
TGGAGGAGGCCACCAACCACCTGTTCTGCCAGCGACGCCAACGCCTACGACACCCGAGGT  
GCACAACCTGTGGGCCACCCACGCCCTGCGTGCCTGCCACCGACCCAAACCCCAAGGGAGATCGTGT  
GGAGAACGTGACCGAGAACTCAACATGTGGAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGGCCGTGAAGGCCCTGCGTGAAGCTGACCCCCCTGCGTGT  
GGGGCAGGGAACTGCAACACCAGCGTGTACCCAGGCCCTGCCAACGGTGAAGCTGAGCTCGAGCC  
CATCCCCATCCACTACTGCCCTGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT  
CAACGGCAGCGGCCCTGCACCAACGTGAGCAGCGTGCAGTGACCCACGGCATCCGCCCG  
TGGTGGAGCACCCAGCTGCTGTGAACGGCAGGCCCTGCCAGGGAGGGCTGGTGTACCGCAGC  
GAGAACCTCACCGACAACGCCAACGACCATCATCGTCAGCTGAAGGAGAGCGTGGAGATCAA  
CTGCACCCGCCAACAAACACCCGAAGAGCATCACCATCGGCCCCGGCGCCCTTCTA  
CGCCACCGCGACATCATCGCGACATCCGCCAGGCCACTGCAACATCAGCGCGAGAAGT  
GGAACAAACACCCCTGAAGCAGATCGTGAACAGCTGCAGGCCAGTTGGCAACAAGACCATC  
GTGTTCAAGCAGAGCAGCGGGCGACCCGAGATCGTGTACAGCTTCAACTGCCCG  
CGAGTTCTTCACTGCAACAGCACCCAGCTGTTAACAGCACCTGGAACAAACACCATGCC  
CAACAAACACCAACGGCACCATCACCTGCCCTGCCCATCAAGCAGATCATCAACCGCG  
GGCGCAAGGCCATGTACGCCCTGCCAGATCGCTGAGCAGCAACATCACC  
GCCCTGCTGCTGACCCCGCAAGGGAGATCAGGAAACACCCAGGAGATCTTCCGCC  
CGGGGGCGGCACATGCGCAACACTGGCGAGCGAGCTGTACAAGTACAAGGTGGTGAAG  
ATCGAGCCCCCTGGCGTGGCCCCCACCAAGGCCAACGCCGCTGGTGCAGCGCGAGAACG  
CGCCGTGACCCCTGGCGCCATGTTCTGGCTTCTGGGCCGCCGGCAGCACCATGGCGC  
CCGCAGCCTGACCCGTGACCGTGAGGCCGCCAGCTGCTGAGCGGCATCGTCAGCAGCAGA  
ACAACCTGCTGCGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCGTGTGGGCATC  
AAGCAGCTGCAGGCCCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGG  
CATCTGGGCTGCAAGCGCAAGCTGATCTGACCCACCGCCGTGCCCTGGAACGCCAGCTGGA  
GCAACAAAGAGCCTGGACCAGATCTGGAACACATGACCTGGATGGAGTGGAGCGCGAGATC  
GACAACATACCAACCTGATCTACACCCCTGATCGAGGAGAGGCCAGAACCCAGCAGGAGAAAGAA  
CGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTTGAACCTGGTGCACATCAGCA  
AGTGGCTGTTGACATCAAGATCTTCACTGATGTCAGCTGGCGGCCCTGGTGGGCCCTGCGCATCG  
TGTTCACCGTGTGAGCATCGTAACCGCGTGCAGGCCAGGGCTACAGCCCCCTGAGCTTCCAGA  
CCCGCTTCCCGCCCCCGCGGCCCGACGCCCGAGGGCATCGAGGAGGAGGGCGCGAG  
CGCGACCGCGACCGCAGCACCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTG  
CGCAGCCTGTCCTGTTCACTGACCCACCGCCGTGCGCAGCTGATCTGATCGCCGCCGCATC  
GTGGAGCTGCTGGGCCCGCTGGAGGCCCTGAAAGTACTGGGCAACCTGCTGCAAGTA  
CTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATGCCATGCCGTGG  
CCGAGGGCACCGACCGCATCGAGGTGGCCAGCGCATCGGCCGCCCTGCTGTAACCTCGAG  
CCCGCCGCATCCGCCAGGGCTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 29

## SEQUENCE LISTING

<110> Chiron Corporation

<120> MODIFIED HIV ENV POLYPEPTIDES

<130> 1605.100

<140>

<141>

<160> 26

<170> PatentIn Ver. 2.0

<210> 1

<211> 856

<212> PRT

<213> Human immunodeficiency virus

<400> 1

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg  
1 5 10 15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu  
20 25 30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala  
35 40 45

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu  
50 55 60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn  
65 70 75 80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp  
85 90 95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp  
100 105 110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser  
115 120 125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser  
130 135 140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn  
145 150 155 160

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe  
165 170 175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys  
180 185 190

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val  
 195 200 205

Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala  
 210 215 220

Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr  
 225 230 235 240

Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser  
 245 250 255

Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Val Val Ile  
 260 265 270

Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu  
 275 280 285

Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg  
 290 295 300

Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile  
 305 310 315 320

Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala  
 325 330 335

Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln  
 340 345 350

Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp  
 355 360 365

Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Glu Phe Phe Tyr  
 370 375 380

Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp  
 385 390 395 400

Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu  
 405 410 415

Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys  
 420 425 430

Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn  
 435 440 445

Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu  
 450 455 460

Ser Glu Ile Phe Arg Pro Gly Gly Asp Met Arg Asp Asn Trp Arg  
 465 470 475 480

Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val  
 485 490 495

Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala  
 500 505 510

Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser  
 515 520 525  
  
 Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu  
 530 535 540  
  
 Leu Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu  
 545 550 555 560  
  
 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu  
 565 570 575  
  
 Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu  
 580 585 590  
  
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val  
 595 600 605  
  
 Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn  
 610 615 620  
  
 His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser  
 625 630 635 640  
  
 Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn  
 645 650 655  
  
 Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp  
 660 665 670  
  
 Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile  
 675 680 685  
  
 Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile  
 690 695 700  
  
 Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His  
 705 710 715 720  
  
 Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu  
 725 730 735  
  
 Gly Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser  
 740 745 750  
  
 Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr  
 755 760 765  
  
 His Arg Leu Arg Asp Leu Leu Ile Val Thr Arg Ile Val Glu Leu  
 770 775 780  
  
 Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu  
 785 790 795 800  
  
 Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn  
 805 810 815  
  
 Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val  
 820 825 830

Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg  
 835 840 845

Gln Gly Leu Glu Arg Ile Leu Leu  
 850 855

<210> 2  
 <211> 847  
 <212> PRT  
 <213> Human immunodeficiency virus

<400> 2  
 Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Gly  
 1 5 10 15

Gly Thr Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Val Glu Lys  
 20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr  
 35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val  
 50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
 65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
 85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp  
 100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
 115 120 125

His Cys Thr Asn Leu Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp  
 130 135 140

Lys Glu Met Asp Arg Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr  
 145 150 155 160

Thr Ser Ile Arg Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys  
 165 170 175

Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile  
 180 185 190

Asn Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe  
 195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu  
 210 215 220

Lys Cys Asn Asp Lys Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val  
 225 230 235 240

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln  
 245 250 255  
 Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser  
 260 265 270  
 Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu  
 275 280 285  
 Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser  
 290 295 300  
 Ile Thr Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile  
 305 310 315 320  
 Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn  
 325 330 335  
 Asn Thr Leu Lys Gln Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn  
 340 345 350  
 Lys Thr Ile Val Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val  
 355 360 365  
 Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr  
 370 375 380  
 Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr  
 385 390 395 400  
 Asn Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg  
 405 410 415  
 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln  
 420 425 430  
 Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly  
 435 440 445  
 Gly Lys Glu Ile Ser Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly  
 450 455 460  
 Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val  
 465 470 475 480  
 Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val  
 485 490 495  
 Val Gln Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly  
 500 505 510  
 Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu  
 515 520 525  
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn  
 530 535 540  
 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr  
 545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg  
 565 570 575  
  
 Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys  
 580 585 590  
  
 Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys  
 595 600 605  
  
 Ser Leu Asp Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg  
 610 615 620  
  
 Glu Ile Asp Asn Tyr Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser  
 625 630 635 640  
  
 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys  
 645 650 655  
  
 Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr  
 660 665 670  
  
 Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile  
 675 680 685  
  
 Val Phe Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser  
 690 695 700  
  
 Pro Leu Ser Phe Gln Thr Arg Phe Pro Ala Pro Arg Gly Pro Asp Arg  
 705 710 715 720  
  
 Pro Glu Gly Ile Glu Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser  
 725 730 735  
  
 Ser Pro Leu Val His Gly Leu Leu Ala Leu Ile Trp Asp Asp Leu Arg  
 740 745 750  
  
 Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu Ile  
 755 760 765  
  
 Ala Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu  
 770 775 780  
  
 Lys Tyr Trp Gly Asn Leu Leu Gln Tyr Trp Ile Gln Glu Leu Lys Asn  
 785 790 795 800  
  
 Ser Ala Val Ser Leu Phe Asp Ala Ile Ala Ile Ala Val Ala Glu Gly  
 805 810 815  
  
 Thr Asp Arg Ile Ile Glu Val Ala Gln Arg Ile Gly Arg Ala Phe Leu  
 820 825 830  
  
 His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu  
 835 840 845

<210> 3  
 <211> 2310  
 <212> DNA  
 <213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val120-Ala204

&lt;400&gt; 3

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtcttcg tttcgccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360  
 ggcgcctgcc ccaaggtgag cttcgagccc atccccatcc actactgcgc ccccgccggc 420  
 ttgcgcatcc tgaagtgcaa cgacaagaag ttcaacggca gcgccccctg caccacgtg 480  
 agcaccgtgc agtgcacccca cggcataccgc cccgtggta gcacccagct gctgctgaa 540  
 ggcagcctgg ccgaggaggg cgtgggtatc cgacgcgaga acttcacccga caacgccaag 600  
 accatcatcg tgcagctgaa ggagagcgtg gagatcaact gcacccgccc caacaacaac 660  
 acccgcaaga gcatcaccat cggcccccgc cgccgccttct acgcccacccgg cgacatcatc 720  
 ggcgacatcc gccaggccca ctgcaacatc agcggcgaga agtggaaacaa caccctgaa 780  
 cagatcgtga ccaagctgca ggcccagttc ggcaacaaga ccatcgtgtt caagcagagc 840  
 agcggcggcg accccgagat cgtgtatgcac agcttcaact gcgccggcga gttttctac 900  
 tgcaacagca cccagctgtt caacacgacc tggaaacaaca ccacatggccc caacaacacc 960  
 aacggcacca tcaccctgcc ctgcccgcata aagcagatca tcaacccgtg gcaggagggt 1020  
 ggcaaggcca tgcacccccc ccccatccgc ggccagatcc gtcgcagcag caacatcacc 1080  
 ggcctgctgc tgaccggcga cggcggcaag gagatcagca acaccacccga gatttccgc 1140  
 cccggcggcg ggcacatgcg cgacaaactgg cgacgcgagc tgcataagta caaggtggtg 1200  
 aagatcggcgc ccctggcgt ggccccccacc aaggccaacg gccgcgtgg gcagcgcgag 1260  
 aagcgcggcg tgaccctggg cccatgttc ctgggcattcc tgggcggccgc cggcagcacc 1320  
 atgggcggccc gcagcctgac ctcgcacccgtg caggcccgcc agctgctgag cggcatacg 1380  
 cagcagcaga acaacctgt gcgcgcaccc gaggcccgac agcacctgt gcagctgacc 1440  
 gtgtggggca tcaaggagct gcaggcccgc gtgcgtggccg tggagcgcata cctgaaggac 1500  
 cagcagctgc tggcgcattc gggctgcagc ggcaagctga tctgcaccac cgccgtgccc 1560  
 tggaaacgcca gctggagcaa caagacgcctg gaccagatct ggaacaacat gacctggatg 1620  
 gagtgggagc gcgagatcga caactacacc aacctgtatc acaccctgtat cgaggagagc 1680  
 cagaaccagc aggagaagaa cgagcaggag ctgcgtggagc tggacaagtgg gcccagcctg 1740  
 tggaaactggc tgcacatcag caagtggctg tggatcatca agatcttcat catgatcg 1800  
 ggcggcctgg tggcgcctgcg catcgatgttcc accgtgtgaa gatcgatgcg 1860  
 cagggctaca gccccctgag cttccagacc cgcttcccccg ccccccggg ccccgaccgc 1920  
 cccgaggggca tcgaggagga gggcggcgag cgacgcacccg acggcagcag cccctgggtg 1980  
 cacggcctgc tggccctgtat ctgggcacgc ctgcgcagcc tgcgcgtgtt cagctaccac 2040  
 cgccctgcgcg acctgtatcc gatgcgcgcg cgacgcgtgg agctgctggg ccgcgcggc 2100  
 tgggaggccc tgaagtactg gggcaacctg ctgcgtact gatccaggaa gctgaagaac 2160  
 agcgccgtga gcctttcga cgcacatgcac atgcgcgtgg ccgaggccac cgaccgcac 2220  
 atcgaggtgg cccagcgcac cggccgcgc ttcctgcaca tcccccggc catccgcac 2280  
 ggcttcgagc ggcgcctgct gtaactcgag 2310

&lt;210&gt; 4

&lt;211&gt; 2316

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val120-Ile201

&lt;400&gt; 4

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtcttcg tttcgccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggcggc 360

atcacccagg cctggcccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420  
 gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480  
 aacgtgagca ccgtcagtg caccacggc atccgccccg tggtgagcac ccagctgctg 540  
 ctgaacggca gcctggccga ggagggcgtg gtatccgca gcgagaactt caccgacaac 600  
 gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaaac 660  
 aacaacaccc gcaagagcat caccatcgcc cccggccgca ccttctacgc caccggcgc 720  
 atcatcgccg acatccgca gccccactgc aacatcagcg gcgagaagtg gaacaacacc 780  
 ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840  
 cagagcagcg gcggcgaccc cgagatcgat atgcacagct tcaactgcgg cggcgagttc 900  
 ttctactgca acagcacccca gctgttcaac agcacctgga acaacaccat cggccccaaac 960  
 aacaccaacg gcaccatcac cctgcccgtc cgcatcaagc agatcatcaa cgcgtggcag 1020  
 gaggtggcca aggccatgta cgccccccccc atccgcggcc agatccgctg cagcagcaac 1080  
 atcaccggcc tgctgctgac ccgcgcacggc ggcaaggaga tcagcaacac caccgagatc 1140  
 ttccgccccg gcggggcga catgcgcgac aactggcga gcgagctgt acaagtacaag 1200  
 gtggtaaga tgcagccccct gggcgtggcc cccaccaagg ccaagcggc cgtggcag 1260  
 cgcgagaagc gcgcgtgac cctggggcgc atgttccctgg gcttccctgg cggccggc 1320  
 agcaccatgg gcgcggcag cctgaccctg accgtgcagg cccgcccagct gctgagcggc 1380  
 atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440  
 ctgaccgtgt gggcatcaa gcagctgcag gcccgcgtgc tggccgtgga ggcgttccctg 1500  
 aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgc 1560  
 gtgcctgca acgcagctg gagcaacaag agcctggacc agatctgaa caacatgacc 1620  
 tggatggagt gggagcgcga gatcacaac tacaccaacc tgatctacac cctgatcgag 1680  
 gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtggcc 1740  
 agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatatg 1800  
 atcgtggccg gcctggggc cctgcgcatac gtgttccaccc tgcgtgagcat cgtgaaccgc 1860  
 gtgcgcagg gctacagccc cctgagcttc cagacccgct tcccccccccc cgcggcccc 1920  
 gaccgccccg agggcatcga ggaggaggc ggcgagcgcg accgcgaccg cagcagcccc 1980  
 ctgggtgcacg gcctgtggc cctgatctgg gacgacctgc gcagcctgtg cctgttccage 2040  
 taccaccggc tgcgcgaccc gatcctgatc gcccggcga tcgtggagct gctggccgc 2100  
 cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160  
 aagaacagcg ccgtgagcct gttcgcacgca atcgcacatcg cctggtggca gggcaccgc 2220  
 cgcacatcg aggtggccca ggcacatcgcc cgcgccttcc tgcacatccc cgcgcacatc 2280  
 cgccaggcgt tcgagcgcgc cctgctgtaa ctgcag 2316

&lt;210&gt; 5

&lt;211&gt; 2322

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val120-Ile201B

&lt;400&gt; 5

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcaatcttcg ttcggccca gcccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccacccctg ttctgcgcga ggcgcacccaa ggccctacgc 180  
 acccgagggtc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaaggccctg cgtgcggc 360  
 atcaccacgg cctggcccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420  
 gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480  
 aacgtgagca ccgtcagtg caccacggc atccgccccg tggtgagcac ccagctgctg 540  
 ctgaacggca gcctggccga ggagggcgtg gtatccgca gcgagaactt caccgacaac 600  
 gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaaac 660  
 aacaacaccc gcaagagcat caccatcgcc cccggccgca ccttctacgc caccggcgc 720  
 atcatcgccg acatccgca gccccactgc aacatcagcg gcgagaagtg gaacaacacc 780  
 ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840  
 cagagcagcg gcggcgaccc cgagatcgat atgcacagct tcaactgcgg cggcgagttc 900  
 ttctactgca acagcacccca gctgttcaac agcacctgga acaacaccat cggccccaaac 960

aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020  
 gaggtggca aggccatgta cgcccccccc atcccgccgca agatccgctg cagcagcaac 1080  
 atcaccggcc tgctgctgac cccgcacggc ggcaaggaga tcagcaacac caccgagatc 1140  
 ttccggcccg gggccggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200  
 gtggtaaga tcgagccct gggcggtggc cccaccaagg ccaagcgcgg cgtggctgag 1260  
 cccgagaage gcgcgtgac cctggggccg atgttctgg gtttcctgg cggccggc 1320  
 agcaccatgg gcgcggcag cctgaccctg accgtgcagg cccgcagct gctgagcggc 1380  
 atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgag 1440  
 ctgaccgtgt gggcatcaa gcagctgca gccccgtgc tggccgtgga ggcgtacctg 1500  
 aaggaccagg agctgctggg catctggggc tgcagcggca agctgatctg caccaccgca 1560  
 gtgcctggg acgcaagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620  
 tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680  
 gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtggcc 1740  
 agcctgtggg actgggtcgat catcgcaag tggctgtggt acatcaagat cttcatcatg 1800  
 atcgtggcg gcctgggtgg cctgcgcate gtgttccacg tgctgagcat cgtgaaccgc 1860  
 gtgcgcagg gctacagccc cctgagcttc cagacccgct tccccgcggcc cgcggccccc 1920  
 gacccggcccg agggcatcga ggaggagggc ggccgcgcg accgcgcaccc cagcagcccc 1980  
 ctgggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040  
 taccacccgc tgcgcaccc gatctgtatc gccgcggcga tcgtggagct gctggccgc 2100  
 cgcggctggg aggcctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160  
 aagaacagcg ccgtgagccct gttcgacgccc atcgccatcg ccgtggccga gggcaccgac 2220  
 cgcacatcg aggtggccca ggcacatcgcc cgcgccttcc tgcacatccc cgcgcgcate 2280  
 cggcaggcgt tcgagcgcgc cctgctgtaa ctcgagcgtg ct 2322

&lt;210&gt; 6

&lt;211&gt; 2328

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Lys121-Val200

&lt;400&gt; 6

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtctcg ttccggccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtggg aggagccac caccacccctg ttctgcgcca gcgcacgcggaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccacccgaccc caaccccgag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360  
 cccgtgtatca cccagccctg ccccaaggtg agcttcgagc ccatacccat ccactactgc 420  
 gccccggccg gttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcggccccc 480  
 tgcaccaacg tgagcaccgt gcaatgcacc cacggcatcc gccccgtgtt gacccacccag 540  
 ctgctgctga acggcgcct ggccgaggag ggcgtggta tccgcagcga gaacttcacc 600  
 gacaacgcaca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660  
 cccaacaaca acacccgcac gacatcacc atcgcccccg gccgcgcctt ctacgccacc 720  
 ggcgacatca tcggcgacat cccgcaggcc cactgcaaca tcagcggcga gaagtggaa 780  
 aacacccctga agcagatcgatc gaccaagctg caggccctgt tcggcaacaa gaccatcg 840  
 ttcaaggcaga gcagcggcgg cgaccccgag atcgtgtatgc acagcttcaa ctgcggcggc 900  
 gagttcttct actgcaacacg caccctgttgc acatcgatcc cctggaaacaa caccatcg 960  
 cccaacaaca ccaacggcac catcaccctg ccctgcgcga tcaagcagat catcaaccgc 1020  
 tggcaggagg tggcagggc catgtacgccc ccccccattcc gcggccagat cgcgtgcagc 1080  
 agcaacatca cccgcctgtc gtcgacccgc gacggcggca aggatcgatc caacaccacc 1140  
 gagatcttcc gccccggcgg cggcgacatg cgcgcacact ggcgcagcga gctgtacaag 1200  
 tacaagggtgg tgaagatcgatc gcccctgggg gtggccccc ccaaggccaa ggcgcgcgtg 1260  
 gtgcagcgcg agaagcgcgc cgtgaccctg ggcgcctatgt tccctgggctt cctggcgcc 1320  
 gcccggcggca ccatggcgcgc cccgcacccctg accctgaccg tgcaggcccg ccagctgctg 1380  
 agcggcgcac tgcagcgcac gaccaacctg ctgcgcgcac tgcaggccca gacgcacccctg 1440  
 ctgcagctgatc ccgtgtgggg catcaaggcag ctgcaggccca ggcgtgtggc cgtgagcgc 1500  
 tacctgaagg accagcagct gctggcgcate tggggctgca gccggcaagct gatctgcacc 1560

accggccgtgc cctggAACgc cagctggAGC aacaAGAGCC tggaccAGAT ctggAACAAc 1620  
 atgacCTTggA tggAGTggA gCGCgAGATC gacaACTACA ccaACCTGAT ctacACCCtG 1680  
 atcgaggAGA gCCAGAAccA gcaggAGAAG aacgAGCAGG agCTGCTggA gCTggACAAg 1740  
 tggccAGC C TGTggAACTg gttcgACATC agCAAGTggC tgggtacAT caAGATCTC 1800  
 atcatGATCg tgggCggCCT ggtgggCCTg cgcatCgtGT tcaccGtGtC gAGCATGtG 1860  
 aaccGCGTgc gCCAGGGCTA cAGCCCCtG agCTtCCAGA cCCGtTCCC CGCCCCCgC 1920  
 ggCCCCgACC gCCCCgAGGG catcgaggAG gagggCggCg agCGCgACCC CGACCCgAGC 1980  
 agCCCCtGG tgcACGGCCT gCTggCCtG atCTGGACg acCTGCGAG cCTGtGcCTG 2040  
 ttcAGCTACC accGCTGCG cgACCTGATC ctGATCggC cCCGcATCtG ggAGCTGCTG 2100  
 gGCCGCGEG gCTggGAGGc cCTGAAGTAC tggggCAACC tGtGcAGTA ctggatCCAG 2160  
 gagCTGAAGA acAGCgCCGT gAGCCTGtC gACGcATCg cCATCgCCGT ggCCGAGGGC 2220  
 accgACCGCA tCATCGAGGT ggCCCAgCg atCGGCCGCG CTTCCtGCA catCCCCCgC 2280  
 cgcatCCGCC agggCTtCGA gCGCgCCtG ctGtAactCG agCGtGtC 2328

&lt;210&gt; 7

&lt;211&gt; 2334

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Leu122-Ser199

&lt;400&gt; 7

gaattcgCCA ccatggatGC aatGAAGAGA gggCTCTGtC gtgtGtGtGtC gCTgtGtGGA 60  
 gcAGtCTtCG tttcgCCAG cggcgtggAG aagCTgtggG tgaccGtGtA ctacggcgtG 120  
 cccgtgtGGA aggaggCCAC caccACCCtG ttctGCGCCA gCGACGCCAA ggcctacGAC 180  
 accgaggGTGc acaacGtGtG gggcACCCAC gCtGCGtGc ccACCGACCC caACCCCCAG 240  
 gagatGtGc tggagaACGT gaccGAGAAC ttcaACATGT ggaAGAAACAA catGGtGGAG 300  
 cagatGcACG aggACATCAT cAGCCTGtG gaccAGAGC tGAAGCCtG cgtGAAGCTG 360  
 ggcAACAGCG tGATCACCCCA ggcctGCCCC aaggTgAGtC tCGAGCCCAT cCCCATCCAC 420  
 tactGCGCC CCGCCGGtT CGCCATCCTG aagtGCAACG acaAGAAAGtT caACGGCAGC 480  
 gGCCCTGCA ccaACGtGAG caccGtGcAG tGcACCCACG gcatCCGCCCC cgtGGtGAGC 540  
 acccAGCTGc tGtGtGACCG cAGCCTGGCC gaggAGGGCg tGGtGatCCG cAGCgAGAAC 600  
 ttcACCGACA acGCCAAGAC catCATGtG cAGtGtGAAG agAGCtGtGGA gatCAACTGc 660  
 acccGCCCCA acaACAACAC ccGCAAGAGC atCACCAtCG gCCCCGGCCG CGCCtTtAC 720  
 gCCACCGGCG acATCATCGG cGACATCCG cAGGCCACT gCAACAtCG cGGCGAGAAg 780  
 tggAACAAcA cCtGtGAAGCA gatCGtGACC aAGtGtGAGG cCCAGtTCG caACAAGACC 840  
 atCGtGtTCA agCAGAGCAG CGCGGGCGAC cCCAGAGAtCg tGatGcACAG cttCAACTGc 900  
 gGCAGCAGtT ctttCTACTG caACAGCACC cAGtGtTCA acAGCACtGtG gaACAACACC 960  
 atCGGCCCCA acaACACCAA CGCACCAtC ACCtGtCCtG gCCGtCATCAA gCAGAtCATC 1020  
 aaccGtGGC aggAGGTGGG caAGGCCAtG tACGCCtCC CCAtCCGCGG CCAGAtCCGc 1080  
 tGcAGCAGCA acATCACCGG CctGtGtGtG ACCCGCgACG gCGGCAAGGA gatCAGCAAC 1140  
 accACCGAGA tCTtCCGCCCC CGGCGGGCGC gACAtGCGtG ACCtGtGGCG cAGCAGtGtG 1200  
 tacaAGtTACA agGTGGtGAA gatCGAGGCC CTTGGGtGtG CCCCCACCAA gGCCAAGCGC 1260  
 CGCGtGGtGC AGCGCAGAA gCGCgCCGtG ACCtGtGGCG CCAtGtCCtG gGGtCCtG 1320  
 gGCAGCAGtT gCAGCACCAT gGGCGCCCGC AGCtGACCC tGaccGtGCA gGCCCGCCAG 1380  
 ctGtGAGCG gcatCGtGCA gCAGCAGAAc AACtGtGtGc gCCtGAtCGA gGCCCAAGCAG 1440  
 cacCTGtGc AGtGACCGT GTGGGtGAtC AAGCAGtGc AGGCCCCGtG gCTGGCCtG 1500  
 gAGCtGtAcc tGAAGGACCA gCAGtGtGtG gGAtCtGtGG gCTGcAGCGG caAGtGAtC 1560  
 tGcACCACCG CGtGtCCtG gAAGCtGtGtG tGGAGCAACA AGAGCtGtGtG CCAGAtCTtG 1620  
 aacaACAtGA cCTGGAtGGA GTGGGAGCGC gAGAtGtGACA AtTACACCAA CctGAtCTAC 1680  
 accCtGAtGtG AGGAGAGCCA gAACCCAGAG gAGAAGAAcG AGCAGGAGtC gCTGGAGtG 1740  
 gacaAGtGGG CCAGCtGtG gAActGtGtC gACAtCAGCA AGtGGtGtGtG tGAtCATCAAG 1800  
 atCtTCAtCA tGAtGtGtGGG CGGCTGGtG gGCtGCGtA tGtGtGtCAG cGtGtGAGC 1860  
 atCGtGAACc gCGtGCGCA gGGtACAGC CCCCtGAGtC tCCAGACCCG CTTCCCCGtC 1920  
 CCCCCGGCC CCGACCCGCCCC CGAGGGtGAtC gAGGAGGAGG gCGGCGAGCG CGACCCGtC 1980  
 CGCAGCAGCC CCtGtGtCA CGGCTGtGtG gCCtGAtCtC gGAGCAGAtC gCGCAGCtG 2040  
 tGcCTGtTCA gCTACCAACG CCTGCGCGAC ctGAtCtGtA tGtGtGtGtGtGtGtG 2100  
 ctGtGtGGCC gCCGCGGtGtG gGAGGtCCtG aAGtACTGGG gCAACtGtGtC gCAGtACTGG 2160

atccaggagc tgaagaacag cgccgtgagc ctgttcgacg ccatgccat cgccgtggcc 2220  
 gagggcaccg accgcacatcat cgaggtggcc cagcgcatcg gcccgcctt cctgcacatc 2280  
 ccccgccgca tccgcccaggc ttccgagcgc gcccgtctgt aactcgagcg tgct 2334

&lt;210&gt; 8

&lt;211&gt; 2316

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val120-Thr202

&lt;400&gt; 8

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgtctgt gctgtgtgga 60  
 gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcg 120  
 cccgtgtgga aggaggccac caccacccctg ttctgcgcca ggcacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggggggc 360  
 gcccacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420  
 gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480  
 aacgtgagca cctgtcgatg caccacacggc atccgcggc tggtgagcac ccagctgtg 540  
 ctgaacggca gcctggccga ggagggcg 600  
 gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac cgcggccaaac 660  
 aacaacaccc gcaagagcat caccatcgcc cccggccg 720  
 atcatcgccg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780  
 ctgaagcaga tcgtgaccaa gctgcaggcc cagttcg 840  
 cagagcagcg gcggcgaccc cgagatcg 900  
 ttctactgca acagcaccca gctgttcaac agcacctgga acaacacccat cggggccaaac 960  
 aacaccaacg gcaccatcac cctgcctgc cgcatca 1020  
 gaggtgggca aggccatgt 1080  
 atcaccggcc tgctgtgac cgcgcacggc ggcaaggaga 1140  
 ttccggcccg gcggcg 1200  
 gtggtaaga tcgagccct 1260  
 cgcgagaagc gcgcgtgac cctggggc 1320  
 agcaccatgg gcgcggcag cctgaccctg accgtgcagg cccgc 1380  
 atcgtgcagc agcagaacaa cctgctgc 1440  
 ctgaccgtgt gggcatcaa gcagctgc 1500  
 aaggaccagc agctgctgg 1560  
 gtgcctgga acgcccagctg gagcaacaag 1620  
 tggatggagt gggagcgc 1680  
 gagagccaga accagcagga 1740  
 agcctgtgga actggttcg 1800  
 atcgtggccgc gcctgggtgg 1860  
 gtgcgc 1920  
 gaccgc 1980  
 ctgggtgcacg gcctgtggc 2040  
 taccaccgccc tgcgcgac 2100  
 cgcggctggg aggcctgaa 2160  
 aagaacagcg ccgtgagcct 2220  
 cgcacatcg aggtggccca 2280  
 cgccagg 2316  
 ggcgcgcgc cctgctgtaa ctgcag

&lt;210&gt; 9

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Description of Artificial Sequence: Trp427-Gly431

&lt;400&gt; 9

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtctcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca ggcacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcggtc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 acccccccgt gcgtgaccct gcactgcacc aacctaaga aacccaccaa caccaagagc 420  
 agcaactgga aggagatgga cccggcggag atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgca acaagatgca gaaggaggtac gcccgttct acaagctgga cgtggtgc 540  
 atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccac 600  
 gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgc 540  
 gcccggcttc 660  
 gccatctgaa agtgcacgca caagaagttc aacgcagcg gcccgtgcac caacgtgagc 720  
 accgtgcagt gcacccacgg catccggccc gtggtagca cccagctgct gctgaacggc 780  
 agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaaagacc 840  
 atcatcggtc agctgaagga gagcgtggag atcaactgca cccggcccaa caacaacacc 900  
 cgcaagagca tcaccatcg 960  
 gacatccgccc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcggtgacca agctgcaggc ccagttcg 1080  
 ggcggcgacc ccgagatcgt gatgcacagc ttcaactg 1140  
 aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggcccaa caacaccaac 1200  
 ggcaccatca ccctgccc 1260  
 atgtacgccc ccccatccg cggccagatc cgctgcagca gcaacatc 1320  
 ctgaccggcg acggcg 1380  
 ggcgacatgc ggcgacaactg ggcgac 1440  
 cccctggcg tggcccccac 1500  
 gtgaccctgg ggc 1560  
 cgcagctga ccctgaccgt gca 1620  
 aacaacctgc tgcgc 1680  
 atcaagcagc tgcaggccc 1740  
 ctgggc 1800  
 agctggagca acaagac 1860  
 cgcgagatcg acaactacac 1920  
 caggagaaga acgagcagga 1980  
 ttcgacatca gca 2040  
 gtgggc 2100  
 agccccc 2160  
 atcgaggagg agggcg 2220  
 ctggcc 2280  
 gacctgatcc tgcgc 2340  
 ctgaagtact gggca 2400  
 agcctgttgc acgc 2460  
 gcccagcgca tgcgc 2520  
 cgc 2541

&lt;210&gt; 10

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Arg426-Gly431

&lt;400&gt; 10

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtctcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca ggcacgccaa ggcctacgac 180

accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcggtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 accccctgt gcgtgaccct gcactgcacc aacctgaaga acgcccacca caccaggagc 420  
 agcaactgga aggagatgga cgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgc 540  
 atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccagg 600  
 gcctgccccca aggtgagctt cgagccatc cccatccact actgcgc 660  
 gccatcctga agtgcacacg aagaaggatc aacggcagcg gcccctgcac caacgtgagc 720  
 accgtgcagt gcacccacgg catccgc 780  
 agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaaagacc 840  
 atcatcggtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900  
 cgcaagagca tcaccatcg 960  
 gacatccgccc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcggtgacca agctgcaggc ccagttcg 1080  
 ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gccggcgagg 1140  
 aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200  
 ggcaccatca ccctgccc 1260  
 atgtacgccc ccccatccg cggccagatc cgctgcagca gcaacatc 1320  
 ctgaccggcg acggcgccaa ggagatcgc aacaccaccc agatcttccg ccccgccg 1380  
 ggcgacatgc ggcacaactg ggcgacg 1440  
 cccctggccg tggcccccac caaggccaag cggccgtgg tgcagcgcga gaagcgcg 1500  
 gtgaccctgg ggcacatgtt cctggcttc ctggcgccg cggcagcac 1560  
 cgcagcctga ccctgaccgt gcaaggcc 1620  
 aacaacctgc tgcgc 1680  
 atcaagcagc tgcaggccc 1740  
 ctgggcacatc ggggcacatc 1800  
 agctggagca acaagac 1860  
 cgcgacatc acaactacac caac 1920  
 caggagaaga acgac 1980  
 ttcgacatca gcaagtggct tggatcatc aagatcttca tcatgatc 2040  
 gtgggcctgc gcatcg 2100  
 agccccc 2160  
 atcgaggagg aggccg 2220  
 ctggccctga tctggacga ctcgcgc 2280  
 gacctgatcc tgcgc 2340  
 ctgaagta 2400  
 agccgttcg acgc 2460  
 gcccagcgca tggccgc 2520  
 cgcgcctgc tgtaactcg 2541

&lt;210&gt; 11

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Arg426-Gly431B

&lt;400&gt; 11

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgg 60  
 gcagtcttcg tttcgcc 120  
 cccgtgtgg 180  
 accgaggtgc acaacgtgtg gcccacccac gcctgcgtgc ccaccgaccc 240  
 gagatcggtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 accccctgt gcgtgaccct gcactgcacc aacctgaaga acgcccacca caccaggagc 420  
 agcaactgga aggagatgga cgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgc 540

atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600  
 gcctgccccca aggtgagctt cgagccatc cccatccact actgcgc(ccc cgccggcttc 660  
 gccatcctga agtgcacacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720  
 accgtgcagt gcaccacacgg catccgcccc gtggtgagca cccagctgt gctgaacggc 780  
 agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840  
 atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900  
 cgcaagagca tcaccatcgg ccccgccgc gccttctacg ccaccggcga catcatcgcc 960  
 gacatccgccc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcggtgacca agctgcaggc ccagttcgcc aacaagacca tcgtgttcaa gcagagcagc 1080  
 ggccggcgacc ccgagatcgt gatgcacacg ttcaactgcg gccggcgagg tttctactgc 1140  
 aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200  
 ggcaccatca ccctgccccctg cccgcacatca cagatcatca accgcggcag cgccaaaggcc 1260  
 atgtacgccc ccccccattccg cggccagatc cgcgcacatca gcaacatcac cggcctgtcg 1320  
 ctgaccgcgc acggccgcaa ggagatcagc aacaccaccc agatcttccg ccccgccgc 1380  
 ggcgcacatgc ggcacacactg ggcgcacatgc ctgtacaagt acaaggtggt gaagatcag 1440  
 cccctggccg tggcccccac caaggccaag cggccgtgg tgcagcgcga gaagcgcgc 1500  
 gtgaccctgg ggcacatgtt cctggccttc ctggcgcgc cccgcacatgc catggggcc 1560  
 cgcacgcctga ccctgaccgt gcaggccccgc cagctgcac ggcgcacatgc gcaacgcac 1620  
 aacaacactgc tgcgcgcacatcg cagggccac cagcacatgc tgcagcgtgac cgtgtgggc 1680  
 atcaagcagc tgcaggccccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740  
 ctggccatct ggggctgcag cggcaagctg atctgcacca cccgcgtgccc ctggaaacgc 1800  
 agctggagca acaagaccc gacccatgc tggacaacaata tgacctggat ggagtggag 1860  
 cgcgcacatgc acaactacac caacactgatc tacaccctga tcgaggagag ccagaacccag 1920  
 caggagaaga acgacgacca gctgctggag ctggacaagt gggccagccct gtggaaactgg 1980  
 ttcgacatca gcaagtggct gtggatccatc aagatcttca tcatgatcgt gggccgcctg 2040  
 gtggccctgc gcatcgtgtt caccgtgcgt agcatcgtga accgcgtgcg ccagggtctac 2100  
 agcccccctga gcttccagac cccgcgttcccc gccccccgcg gccccgcaccc ccccgaggcc 2160  
 atcgaggagg agggccggca ggcgcacccg gacccgcac gccccctggt gcacggccctg 2220  
 ctggccctga tctggacca cctgcgcacgc ctgtgcctgt tcagctacca cccgcctgcgc 2280  
 gacccatcc tgcgcgcac cccgcacatgc gacccatgc tggacaacaata tgacctggat ggagtggag 2340  
 ctgaagttact ggggcaaccc gctgcagttac tggatccagg agctgaagaa cagcgcgcgtg 2400  
 agccctttcg acgcacatgcg catgcgcgcg gccgaggccca ccgcacccat catcgaggtg 2460  
 gcccacgcgc tggccgcgc cttccctgcac atccccccgc gcatccgcac gggcttcgag 2520  
 cgccgcctgc tgtaactcga g 2541

&lt;210&gt; 12

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Arg426-Lys432

&lt;400&gt; 12

gaattcgcacccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgg 60  
 gcagtcttcg tttcgccccag cggccgtggag aagctgtggg tgaccgtgtt ctacggcgtg 120  
 cccgtgtggag aggaggccac caccacccctg ttctgcgcac ggcacgcacca ggcctacac 180  
 accgagggtgc acaacgtgtg ggcacccac gcctgcgtgc ccaccgcaccc caacccccc 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtgag 300  
 cagatgcacg aggacatcat caccgtgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 acccccccgt gcgtgaccct gcactgcaccc aacctgaaga acgcacccaa caccaccc 420  
 agcaactggagaggatggg cccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgcac acaagatgc aacaggatgtac gcccgttct acaagctggc cgtgggtggcc 540  
 atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600  
 gcctgccccca aggtgagctt cgagccatc cccatccact actgcgc(ccc cgccggcttc 660  
 gccatcctga agtgcacacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720  
 accgtgcagt gcaccacacgg catccgcccc gtggtgagca cccagctgt gctgaacggc 780  
 agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840  
 atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900

cgcaagagca tcaccatcg ccccgccgc gccttctacg ccaccggcga catcatcg 960  
 gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcgtgacca agctgcaggc ccagttcg aacaagacca tcgtgttcaa gcagagcagc 1080  
 ggcggcgacc ccgagatcgt gatgcacagc ttcaactcg cgccgcgagg tttctactgc 1140  
 aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggcccaa caacaccaac 1200  
 ggcaccatca ccctgcccgt ccgcacatcaag cagatcatca accgcggcgg caacaaggcc 1260  
 atgtacgccc ccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320  
 ctgaccgcg acggcggcaa ggagatcagc aacaccaccc agatcttcg cccggcggc 1380  
 ggcgacatgc ggcacaactg ggcgcgagc ctgtacaagt acaaggttgt gaagatcgag 1440  
 cccctggcg tggcccaaag cagggccaa cgccgcgtgg tgcagcgcga gaagcgcgc 1500  
 gtgaccctgg ggcceatgtt cctggcttc ctggcgccg cggcgcgac catggcgcc 1560  
 cgcagcctga ccctgaccgt gcaaggccgc cagctgctga ggcgcacatcg gcagcagcag 1620  
 aacaacctgc tgcgcgcatt cggggccag cagcacatcg tgcagctgac cgtgtgggc 1680  
 atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740  
 ctgggcacatcg ggggctgcag cggcaagctg atctgcacca cggccgtgccc ctggAACGCC 1800  
 agctggagca acaagagcct gacccagatc tggacaaca tgcacccgtt ggagtggag 1860  
 cgcgagatcg acaactacac caacctgatc tacaccctga tgcaggagag ccagaaccag 1920  
 caggagaaga acgagcagga gctgctggag ctggacaaat gggccagcct gtggaaactgg 1980  
 ttgcacatca gcaagtggct gtggatcatc aagatcttca tcatgtatcg gggccgcctg 2040  
 gtgggcctgc gcatcggtt caccgtgtc agcatcgta accgcgtgcg ccaggcgtac 2100  
 agccccctga gcttccagac cccgtcccccc gccccccgcg gccccgaccg ccccgagggc 2160  
 atcgaggagg agggcggcga ggcgcaccgc gaccgcagca gccccctggt gcacggcctg 2220  
 ctggccctga tctggacga cctgcgcagc ctgtgcctgt tcaagctacca cgcctgcgc 2280  
 gacctgatcc tgcgcgcgc cgcacatcg tggatccagg agctgaagaa cagcgcctg 2340  
 ctgaagtact gggcaacact gctgcagttac tggatccagg agctgaagaa cagcgcctg 2400  
 agcctgttcg acgcacatcg catgcgcgtg gccgaggcga ccgaccgcacatcgaggtg 2460  
 gcccagcgcga tcggccgcgc ctccctgcac atccccccgc gcatccgcac gggcttcgag 2520  
 cgcgcctgc tgtaactcg 9 2541

&lt;210&gt; 13

&lt;211&gt; 2535

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Asn425-Lys432

&lt;400&gt; 13

gaattcgcaca ccatggatgc aatgaagaga gggctctgct gtgtgtgt gctgtgtgga 60  
 gcagtcttcg ttgcggccag cgcgcgtggag aagctgtggg tgaccgtgtt ctacggcg 120  
 cccgtgtggg aggaggccac caccacccctg ttctgcgcac ggcacgcacca ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgcaccc caaccccccag 240  
 gagatcgatc tggagaacact gaccgagaac ttcaacatgt ggaagaacaa catggtgag 300  
 cagatgcacg aggacatcat cagcgtgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 accccccctgt gcgtgaccct gcaactgcaccc aacctgaaga acgcacccaa caccacggc 420  
 agcaactggg aggagatggg cgcggcgag atcaagaact gcaacttca ggtgaccacc 480  
 agcatccgc acaagatgca gaaggagttac gcccgttct acaagctgga cgtgggtccc 540  
 atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccac 600  
 gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgcctt cgcggctt 660  
 gccatcctga agtgcacacgca caagaagttc aacggcagcg gcccctgcac caacgtgac 720  
 accgtgcagt gcacccacgg catccgcctt gttgtggca cccagctgt gctgaacggc 780  
 agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgcacca cgccaagacc 840  
 atcatcgatc agctgaaggg gacgcgtggag atcaactgca cccggcccaa caacaacacc 900  
 cgcaagagca tcaccatcg ccccgccgc gccttctacg ccaccggcga catcatcg 960  
 gacatccgc aaggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcgtgacca agctgcaggc ccagttcg aacaagacca tcgtgttcaa gcagagcagc 1080  
 ggcggcgacc cccgagatcgt gatgcacagc ttcaactcg cgccgcgagg tttctactgc 1140  
 aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggcccaa caacaccaac 1200  
 ggcaccatca ccctgcccgt ccgcacatcaag cagatcatca acggcccaa ggccatgtac 1260

gcccccccca tccgcggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc 1320  
 cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgccccgg cgccggcgac 1380  
 atgcgcgaca actggcgcag cgagctgtac aagtacaagg tggtaagat cgagccccctg 1440  
 ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgccgtgacc 1500  
 ctggcgcca tggctctggg cttcctggc gccgccccca gcaccatggg cgcccgccagc 1560  
 ctgaccctga ccgtgcaggg cccgcagctg ctgagcggca tcgtgcagca gcagaacaac 1620  
 ctgctgcgcg ccacgcggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1680  
 cagctgcagg cccgcgtgct ggccgtggag cgctacactga aggaccagca gctgctggc 1740  
 atctggggct gcagcgccaa gctgatctgc accaccgcgg tgccctggaa cgccagctgg 1800  
 agcaacaaga gcctggacca gatctggAAC aacatgacct ggtggagtg ggagcgcgag 1860  
 atcgacaact acaccaacact gatctacacc ctgatcgagg agagccagaa ccagcaggag 1920  
 aagaacgagc aggagctgct ggagctggac aagtggggca gcctgtggaa ctgggtcgac 1980  
 atcagcaagt ggctgtggta catcaagatc ttcatcatga tcgtggggcg cctgggtggc 2040  
 ctgcgcatcg tggctaccgt gctgagcatc gtgaaccgcg tgccgcaggg ctacagcccc 2100  
 ctgagcttc agacccgctt cccgcggggcc cgccggggcc accggggccg gggcatcgag 2160  
 gaggagggcg gcgagcgca cccgcggccg agcagcccccc tgggtgcacgg cctgctggcc 2220  
 ctgatctggg acgacactgcg cccgcgtgc ctgatcgact accaccgcct ggcgcacctg 2280  
 atcctgatcg cccgcggcat cgtggagctg ctggggccggc gccgcgtggaa ggccctgaag 2340  
 taactggggca acctgctgca gtactggatc caggagctga agaacagcgc cgtgagcctg 2400  
 ttgcacgcca tcgcacatcg cgtggccggag ggcaccgacc gcatcatcgaa ggtggcccg 2460  
 cgcacatcgcc ggcgccttc gacatcccc ccgcgcaccc gccagggtt ccgcgcgccc 2520  
 ctgctgtaaac tcgag 2535

&lt;210&gt; 14

&lt;211&gt; 2529

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Ile424-Ala433

&lt;400&gt; 14

gaattcgcca ccatggatgc aatgaagaga gggctctgct tgggtgcgt gctgtgtgg 60  
 gcagtctcg tttcgccag cccgcgtggag aagctgtggg tggaccgtgtt ctacggcggt 120  
 cccgtgtggg aggaggccac caccaccctg ttctgcgc ggcacgcggaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggcacccac gcctgcgtgc ccaccgcggcc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtgag 300  
 cagatgcacg aggacatcat cccgcgtgtt gaccgagggcc tggaccgtt cgtgaagctg 360  
 acccccccgt gcgtgaccct gcactgcacc aacctgaaga acgcacccaa caccacggc 420  
 agcaacttggg aggagatggg cccgcggccgg atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgca acaagatgca gaaggaggtac gccctgttca acaagctggc cgtgtgtccc 540  
 atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagctg gatcaccagg 600  
 gcctgcggccca aggtgagctt cccgcgttcc cccatccact actgcgcggcc cgccggcttc 660  
 gccatctgtg agtgcacacgca caagaagttc aacggcagcg gcccctgcac caacgtggc 720  
 accgtgcagt gacccacgg catccgcggc tgggtgagca cccagctgtt gctgaacggc 780  
 agcctggccg aggagggcg tggatccgc agcgagaact tcaccgcacca cgccaaagacc 840  
 atcatcgtgc agctgaaggaa gagcgtggag atcaactgca cccgcggccaa caacaacacc 900  
 cgcacacggca tcaccatcg cccgcggccgc gccttctacg ccaccggcga catcatcgcc 960  
 gacatccgca aggccactg caacatcgac ggcgagaagtt ggaacaacac cctgaagcag 1020  
 atcggtgacca agctgcaggc ccaggatcgac aacaagacca tcgtgttcaa gcagagcgc 1080  
 ggcggcgacc cccatccgcg cccatccgcg ctgcgcggcgc aacatcaccg gcctgcgtt gatcaccatcg 1140  
 aacagcaccg agctgttcaa cggccatcg aacaacaccg tcggcccaaa caacaccaac 1200  
 ggcaccatca ccctgcgttcc cccatccgcg cccatccgcg aacatcaccg gcctgcgtt gatcaccatcg 1260  
 cccatccgcg gccacatccgcg ctgcgcggcgc aacatcaccg gcctgcgtt gatcaccatcg 1320  
 ggcggcgacc agatcagcaa caccacccggat atcttccgcg ccggccggccg cgacatcgcc 1380  
 gacacactggc gcagcggact gtacaagttt aagggtgtt gatcagcggcc cctggccgtt 1440  
 gccccccacca aggccaaagcg ccggcggtt gacgcggaga agcgcggccgt gaccctggc 1500  
 gccatgttcc tgggttccctt gggcgccggcc ggcacacca tggggccccc cggccctgacc 1560  
 ctgaccgtgc aggcccgccca gctgctgagc ggcacatcgac agcagcagaa caacatcgcc 1620

cgccgcacatcg aggcccagca gcacctgctg cagctgaccg tgggggcat caagcagctg 1680  
 caggcccgcg tgctggccgt ggagcgctac ctgaaggacc agcagctgct gggcatctgg 1740  
 ggctgcagcg gcaagctgat ctgcaccacc gccgtgcctt ggaacgcccag ctggagcaac 1800  
 aagagcctgg accagatctg gaacaacatg acctggatgg agtgggagcg cgagatcgac 1860  
 aactacacca acctgatcta caccctgatc gaggagagcc agaaccagca ggagaagaac 1920  
 gacgaggagc tgctggagct ggacaagtgg gccagcctgt ggaactgggt cgacatcagc 1980  
 aagtggctgt ggtacatcaa gatcttcattc atgatcgtgg gcggcctggg gggcctgcgc 2040  
 atcggttca cctgtgtgag catcgtaac cgcgtgcgc aggctacag cccctgagc 2100  
 ttccagaccc gttcccccgc ccccccgcgc cccgaccgc cccgagggcat cgaggaggag 2160  
 ggccgcgagc gcgaccgcgca cccgacgcgc cccctgggtgc acggcctgct ggcctgatc 2220  
 tgggacgacc tgccgacgct gtgcctgttc agctaccacc gcctgcgcga cctgatcctg 2280  
 atcgccgccc gcatcgatgg gctgctggc cgcgcggct gggaggccct gaagtactgg 2340  
 ggcaacctgc tgcaactgt gatccaggag ctgaagaaca ggcgcgtgag cctttcgac 2400  
 gccatcgcca tcgcccgtgc cgagggcacc gaccgcattc tcgaggtggc ccagcgatc 2460  
 ggccgcgcct tcctgcacat ccccccgcgc atccgcagg gcttcgagcg cgccctgctg 2520  
 taactcgag 2529

&lt;210&gt; 15

&lt;211&gt; 2523

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Ile423-Met434

&lt;400&gt; 15

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgg 60  
 gcagtcttcg tttcgcccaag cgccgtggag aagctgtggg tgaccgtgta ctacggcg 120  
 cccgtgtggg aggaggccac caccacccctg ttctgcgcga ggcacgcacaa ggcctacgc 180  
 accgaggtgc acaacgtgtg ggcacccac gcctgcgtgc ccaccgcaccc caacccccc 240  
 gagatcgatc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtaaagctg 360  
 accccctgt gcgtgaccct gcactgcacc aacctgaaga acgcacccaa caccacgc 420  
 agcaactggg aggagatggc cgcggcgag atcaagaact gcagttcaa ggtgaccacc 480  
 agcatccgca acaagatgcgaa gaggaggatc gccctgttct acaagctggc cgtgggtccc 540  
 atcgacaacg acaacaccacg ctacaagctg atcaactgcgca acaccagcgat gatcaccac 600  
 gcctgccccca aggtgagctt cgagccatc cccatccact actgcgcggcc cgcggcttc 660  
 gccatcctgaa gtcgaaacgcgca aacaaatggc aacggcagcg gcccctgcac caacgtgagc 720  
 accgtgcagt gcacccacgg catccgcggc gtggtgagca cccagctgct gctgaacggc 780  
 agcctggcccg aggaggccgt ggtgatccgc agcgagaact tcaccgacaa cgccaaagacc 840  
 atcatcgatgc agctgaaaggag gacggtggag atcaactgcgca cccggcccaaa caacaacacc 900  
 cgcaagagcgtc tcaccatcgccggccgc gccttctacg ccaccggcgacatcatcgcc 960  
 gacatccgccc agggccactg caacatcgcg ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcgatcgatccgc agctgcggcgc ccagttcgcc aacaagacca tcgatgttcaa gcagagcgc 1080  
 ggcggcgacc cccgagatcgatc gatgcacacgc ttcaactgcg gccggcgagtt cttctactgc 1140  
 aacagcaccc agctgttcaa cagcaccatgg aacaacacca tcggcccaaa caacaccaac 1200  
 ggcaccatca ccctgcctg cccgcattcaag cagatcgccg gcatgtacgc ccccccattc 1260  
 cccggccaga tccgcgtgcag cagcaacatc accggcctgc tgctgaccccg cgacggcgcc 1320  
 aaggagatca gcaacaccac cggatcttc cggccggcg gccggcgacat ggcgacaaac 1380  
 tggcgccgcg agctgtacaa gtacaagggtt gtgaagatcg agccctgggg cgtggccccc 1440  
 accaaggccca agcgccgcgt ggtgcagcgc gagaagcgcg ccgtgaccct gggccatg 1500  
 ttccctgggtt tcctggccgc cccggcgccg accatggggc cccgcacccct gaccctgacc 1560  
 gtgcaggcccg cccagatcgatc gacggcgatc gtgcacgcg agaacaaccc gctgcgcgc 1620  
 atcgaggcccg agcgcaccc gctgcacgc accgtgtggg gcatcaacgc gctgcaggcc 1680  
 cccgtgtgtt ccgtggagcg ctacatcgatc gaccgcgcg tgctgggcat ctggggctgc 1740  
 agcggcaagc tgatcgatccgc caccggccgtg ccctggaaacg ccgtggag caacaagagc 1800  
 ctggaccaga tctggaaacaa catgacccctt atggagttggg agcgcgagat cgacaaactac 1860  
 accaaccctgc tctacaccctt gatcgaggag agccagaacc agcaggagaa gaacgagcag 1920  
 gagctgtgttccg agctggacaa gtggccgcgc ctgtggaaact ggttcgacat cagcaagtgg 1980

ctgtggtaca tcaagatctt catcatgatc gtggcgccc tggtggcct ggcacatgtg 2040  
 ttcaccgtgc tgagcatgt gaaccgcgtg cgccaggct acagccccct gagcttccag 2100  
 acccgcttcc cgcggggggc cggccccgac cgccccgagg gcatcgagga ggagggcgcc 2160  
 gagcgcgacc ggcacccgac cagccccctg gtgcacggcc tgctggcct gatctggac 2220  
 gacctgcgca gcctgtgcct gttcagctac caccgcctgc ggcacccgtat cctgatcgcc 2280  
 gcccgcacg tggagctgct gggccggccg ggctgggagg ccctgaagta ctggggcaac 2340  
 ctgctgcagt actggatcca ggagctgaag aacagcgccg tgagcctgtt cgacgccatc 2400  
 gccatcgccg tggccgaggg caccgaccgc atcatcgagg tggcccgccg catcgccgc 2460  
 gccttcctgc acatcccccg cccatccgc caggcctcg agcgcgcccct gctgtactc 2520  
 gag 2523

&lt;210&gt; 16

&lt;211&gt; 2517

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Gln422-Tyr435

&lt;400&gt; 16

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtcttcg tttcgcccaag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtggg aggaggccac caccacccctg ttctgcgcca ggcacccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtgtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 accccccctgt gcgtgaccct gcactgcacc aacctgaaga acgcccaccaa caccaggagc 420  
 agcaactgga aggagatgga cccggccgag atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgca acaagatgca gaaggagtagc gcccctgtt acaagctgga cgtgtgtccc 540  
 atcgacaacg acaacaccag ctacaagctg atcaactgca acaccaggct gatcaccagg 600  
 gcctggccca aggtgagctt cccatccact actgcgcccc cgcggcttc 660  
 gccatcctga agtgcacgca caagaagttc aacggcagcg gcccctgcac caacgtgagc 720  
 accgtgcagt gcaccacccg catccggccc gtggtgagca cccagctgct gctgaacggc 780  
 agcctggccg aggaggccgt ggtgatccgc agcgagaact tcaccgacaa cgccaaagacc 840  
 atcatcgtgc agctgaagga gagcgtggag atcaactgca cccggcccaa caacaacacc 900  
 cgcaagagca tcaccatcgg ccccgccgc gccttctacg ccaccggcga catcateggc 960  
 gacatccgccc aggccactg caacatcage ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcggtacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080  
 ggccggcgacc cccatccgt gatgcacagc ttcaactgca gcccggcgagg tttctactgc 1140  
 aacagcaccc agctgttcaa cccatccgtt aacaacacca tcggcccaa caacaccaac 1200  
 ggcaccatca ccctggccctg cccatccgtt aacccggcga cccatccgtt 1260  
 cagatccgct gcagcagcaa catccatccgtt cccatccgtt aacccggcga cccatccgtt 1320  
 atcagcaaca ccaccatccgtt ggcggccgc acatccgtt aacccggcga cccatccgtt 1380  
 agcgagctgt acaagttcaa ggtggtgagat atcgatccgtt tggccgtggc cccatccgtt 1440  
 gccaagcgcc gcgtggtgca ggcggccgtt aacccggcga cccatccgtt 1500  
 ggcttcctgg ggcggccgg cccatccgtt ggcggccgtt aacccggcga cccatccgtt 1560  
 gcccggccaggc tgctggccgg catccatccgtt aacccggcga cccatccgtt 1620  
 gcccggccaggc acatccgtt ggcggccgg catccatccgtt aacccggcga cccatccgtt 1680  
 ctggccgtgg agcgatccgtt aacccggcga cccatccgtt 1740  
 aagctgtatccgtt ggcggccgg cccatccgtt aacccggcga cccatccgtt 1800  
 cagatctggca acaacatccgtt ggcggccgg catccatccgtt aacccggcga cccatccgtt 1860  
 ctgatctaca ccctgtatccgtt aacccggcga cccatccgtt 1920  
 ctggagctgg acaagttcaa ggcggccgg catccatccgtt aacccggcga cccatccgtt 1980  
 tacatcaaga tcttcatcat gatccatccgtt ggcggccgg catccatccgtt aacccggcga 2040  
 gtgtggccggcatccgtt ggcggccgg catccatccgtt aacccggcga cccatccgtt 2100  
 ttcccccggccccc cccatccgtt aacccggcga cccatccgtt aacccggcga cccatccgtt 2160  
 gaccggcgacc gcagcagccc cctggtgccac ggcggccggcatccgtt aacccggcga cccatccgtt 2220  
 cgcagccctgt gcctgttccatccatccgtt aacccggcga cccatccgtt aacccggcga cccatccgtt 2280  
 atcggtggagc tgctggccgg cccatccgtt aacccggcga cccatccgtt aacccggcga cccatccgtt 2340

cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catgccatc 2400  
 gccgtggccg agggcaccga ccgcacatc gaggtggccc agcgcacatgg ccgcgccttc 2460  
 ctgcacatcc cccggccat ccggcagggc ttcgagcgcg ccctgctgtta actcgag 2517

<210> 17  
 <211> 2517  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Gln422-Tyr435B

<400> 17  
 gaattcgcacccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgg 60  
 gcagtcttcg tttcgcccaag cgccgtggag aagctgtggg tgaccgtgtta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccaccctg ttctcgccca ggcacgcacca ggcctacgac 180  
 accggagggtc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcggtc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaaggccctg cgtgaagctg 360  
 accccctgt gcgtgaccct gcactgcacc aacctgaaga acggccaccaa caccaagagc 420  
 agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480  
 agcatccgca acaagatgca gaaggagtagc gccctgttct acaagctgga cgtggtgc 540  
 atcgacaaacg acaacaccag ctacaagctg atcaactgca acaccagctg gatcaccac 600  
 gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgc 660  
 gccatcttgc agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720  
 accgtgcagt gcacccacgg catccgcccc gtgtgagca cccagctgct gctgaacggc 780  
 agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840  
 atcatcggtc agctgaaagga gagcgtggag atcaactgca cccgccccca caacaacacc 900  
 cgcaagagca tcaccatcg 660  
 gacatccgccc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020  
 atcggtacca agctgcaggc ccagttcg 660  
 ggcggcgacc cccgagatctg gatgcacagc ttcaactg 660  
 aacagcaccc agctgttcaa cagcacctgg aacaacacca tggggcccaaa caacaccaac 1200  
 ggcaccatca ccctgccccctg ccgcacatcaag cagggccccc acggccccc 660  
 cagatccgct gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaaggag 1320  
 atcagcaaca ccaccgagat cttccgcccc ggcggcgccg acatgcgcga caactggcgc 1380  
 agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tggggcgtggc cccaccaag 1440  
 gccaagcgcc gcgtggtgca gcgcgagaag cgcgcgtg 660  
 ggcttcctgg gcgcgcggcgg cagcaccatg ggcgcgc 660  
 gcccgcacgc tgctgagcgg catcgatcg 660  
 gcccagcagc acctgatcg 660  
 ctggccgtgg agcgttaccc 660  
 aagctgatct gcaccaccgc cgtgcctgg 660  
 cagatctgga acaacatgac ctggatggag tggggcgc 660  
 ctgatctaca ccctgatcg 660  
 ctggagctgg acaagtggc cagcctgtgg 660  
 tacatcaaga tcttcatcat gatcg 660  
 gtgctgagca tcgtgaaaccg cgtgcgc 660  
 ttccccgc 660  
 gaccgcgacc gcagcagccc cctggc 660  
 cgcagcctgt gcctgttc 660  
 atcggtggagc tgctggccg 660  
 cagatctgga tccaggagct gaagaacagc 660  
 ggcgtggccg agggcaccga 660  
 ctgcacatcc cccggccat 660

<210> 18  
 <211> 2322  
 <212> DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Leu122-Ser199;  
Arg426-Gly431

&lt;400&gt; 18

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtctcg tttcgcccgag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggagggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccaag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 ggcaacagcg tgatcaccca gcctgcccc aaggtgagct tcgagcccat ccccatccac 420  
 tactgcgccc cggccggctt cgccatcctg aagtgcacg acaagaaggta caacggcagc 480  
 ggccctgtca ccaacgtgag caccgtgcag tgcaacccacg gcatccccc cgtggtgagc 540  
 acccagctgc tgctgaacgg cagcctggcc gaggaggcg tggtgatccg cagcagaac 600  
 ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660  
 acccgccccca acaacaacac ccgcaagagc atcaccatcg gccccggccg cgccttctac 720  
 gcccggccg acatcatcg cagacatccgc caggcccact gcaacatcag cggcagaag 780  
 tggaaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840  
 atcgtgttca agcagagcag cggcggcgcac cccgagatcg tgatgcacag cttcaactgc 900  
 ggcggcgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960  
 atcggccccca acaacaccaa cggcaccatc accctgcctt gccgcataa gcagatcatc 1020  
 aaccgcggcg gggcaaggc catgtacgccc ccccccattcc gggccagat cggctgcagc 1080  
 agcaacatca cggcctgct gctgaccccgac gacggcggca aggagatcag caacaccacc 1140  
 gagatcttcc gccccggcg cggcgacatg cgcgacaact ggcgcagcga gctgtacaag 1200  
 tacaaggtgg tgaagatcga gcccctggc gtggccccc ccaaggccaa ggcggcgtg 1260  
 gtgcagcgcg agaagcgcgc cgtgacccctg ggccatgt tcctgggctt cctggcgcc 1320  
 gcccggcagca ccatgggccc cccgacgcctg accctgaccg tgcaggccc ccaagctgctg 1380  
 agcggcatacg tgcagcagca gaacaacctg ctgcgcgcca tgcaggccca gcagcacctg 1440  
 ctgcagctga cctgtgggg catcaagcag ctgcaggccc gctgtctggc cgtggagcgc 1500  
 tacctgaagg accagcagct gctgggcatac tggggctgca gggcaagct gatctgcacc 1560  
 accggccgtgc cctggaaacgc cagctggagc aacaagagcc tggaccagat ctggaaacaac 1620  
 atgacacttgg tggagtggga ggcgcagatc gacaactaca ccaacctgat ctacaccctg 1680  
 atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgtgga gctggacaag 1740  
 tggggccagcc tgcggaaactg gttcgacatc agcaagtggc tgcggatcatc caagatcttc 1800  
 atcatgatcg tggggccctt ggtggccctg cgcacatgt tcaccgtgt gagcatcg 1860  
 aaccgcgtgc gccaggcgta cagccccctg agcttccaga cccgttccc cccccccgc 1920  
 gggcccccggacc gccccgggg catcgaggag gagggcggcg agcgcgaccc cggccgcagc 1980  
 agccccctgg tgcacccgcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040  
 ttcagctacc accgcctgcg cgcacatgtc ctgatgcgcg cccgcacatgt ggagctgctg 2100  
 gggccggccg gctggggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160  
 gagctgaaga acagcgcgt ggcctgttc gacgcacatcg ccatcgccgt gggccggggc 2220  
 accgaccgcac tcatcgaggat ggcggccgc acgcgcgcgc ctttcctgca catccccccgc 2280  
 cgcacatccgcgcc agggcttcga ggcgcgcgcctg ctgtaaactcg ag 2322

&lt;210&gt; 19

&lt;211&gt; 2322

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Leu122-Ser199;  
Arg426-Lys432

&lt;400&gt; 19

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtctcg tttcgcccgag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120

cccgtgtgga aggaggccac caccacccctg ttctgcgcca ggcacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 ggcaacagcg tgatcaccac ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420  
 tactgcgccc cgcgggctt cgccatcctg aagtgcacg acaagaagtt caacggcage 480  
 ggcccctgca ccaacgtgag caccgtgca gtcacccacg gcatccgccc cgtgtgagc 540  
 acccagctgc tgctgaacgg cagcctggcc gaggaggcg tggtgatccg cagcgagaac 600  
 ttcaccgaca acgccaagac catcatcg cagctgaagg agagcgtgga gatcaactgc 660  
 acccgccccca acaacaacac ccgcaagagc atcaccatcg gccccggccg cgccttctac 720  
 gccaccggcg acatcatcg cagacatccgc cagggccact gcaacatcag cggcgagaag 780  
 tggaaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcg gcaacaagacc 840  
 atcgtttca agcagagcag cggcgccgac cccgagatcg tgatgcacag cttcaactgc 900  
 ggcggcgagt tcttctactg caacagcacc cagctgtca acagcacctg gaacaacacc 960  
 atcggccccca acaacaccaa cggcaccatc accctgcctt gccgcataa gcagatcatc 1020  
 aaccgccccg gcaacaaggc catgtacgcc ccccccattc gcgccagat cgcgtgcagc 1080  
 agcaacatca cccgcctgtc gctgacccgc gacggcgca aggagatcag caacaccacc 1140  
 gagatcttcc gccccggcg cggcgacatg cgcgacact ggccgagcga gctgtacaag 1200  
 tacaagggtgg tgaagatcga gcccctggc gtggccccca ccaaggccaa ggcgcgcgtg 1260  
 gtgcagcgcg agaagcgccg cgtgacccctg ggcgcctatgt tcttgggctt cctggcgcc 1320  
 gccggcagca ccatgggcgc cccgcagctg accctgaccg tgcaggcccg ccagctgctg 1380  
 agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca tgcaggccca gcagcacctg 1440  
 ctgcagctga cccgtgtggg catcaagcag ctgcaggccc gctgtgcgc cgtggagcgc 1500  
 tacctgaagg accagcagct gctgggcattc tgggctgca gcgcaagct gatctgcacc 1560  
 accgcgtgc cctggaaacgc cagctggagc aacaagagcc tggaccagat ctggaaacaac 1620  
 atgacctgga tggagtggg ggcgcagatc gacaactaca ccaacctgtat ctacaccctg 1680  
 atcgaggaga gccagaacca gcaggagaag aacgacgagg agctgtgga gctggacaag 1740  
 tgggcccagcc tggaaactg gttcgacatc agcaagtggc tgggttacat caagatcttc 1800  
 atcatgatcg tggggccctg ggtggccctg cgcacatgt tgcacccgtgt ggcacatgt 1860  
 aaccgcgtgc gccaggggcta cagccccctg agcttccaga cccgcttccc cgcggccgc 1920  
 ggccccggacc gccccggaggg catcgaggag gagggcgccg agcgcgaccg cgcacgcgc 1980  
 agccccctgg tgcacggcct gctggccctg atctggacg acctgcgcag cctgtgcctg 2040  
 ttcagctacc accgcctgcg cgcacatgtc ctgcacccgc cccgcacatgt ggagctgctg 2100  
 ggcgcgcgc gctggggagcc cctgaagttc tggggcaacc tgcgtgcagta ctggatccag 2160  
 gagctgaaga acagcgccgt gacccatgtc gacccatcg ccatcgccgt ggccgagggc 2220  
 accgaccgcgca tcatcgaggt ggcccagcgc atcggccgcg cttccctgca catccccccgc 2280  
 cgcacatcgcc agggcttcga ggcgcggccctg ctgtaaactcg ag 2322

&lt;210&gt; 20

&lt;211&gt; 2322

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Leu122-Ser199;  
Trp427-Gly431

&lt;400&gt; 20

gatttcgcca ccatggatgc aatgaagaga gggctctgtgtgtgtgtgtgtgg 60  
 gcagtcttcg ttctgccttggc cgcgcgtggag aagctgtgg tgaccgtgtatc acggcgctg 120  
 cccgtgtggaa aggaggccac caccacccctg ttctgcgcca ggcacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 ggcaacagcg tgatcaccac ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420  
 tactgcgccc cgcgggctt cgccatcctg aagtgcacg acaagaagtt caacggcage 480  
 ggcccctgca ccaacgtgag caccgtgca gtcacccacg gcatccgccc cgtgtgagc 540  
 acccagctgc tgctgaacgg cagcctggcc gaggaggcg tggtgatccg cagcgagaac 600  
 ttcaccgaca acgccaagac catcatcg cagctgaagg agagcgtgga gatcaactgc 660

acccgccccca acaacaacac ccgcaagagc atcaccatcg gccccggccg cgcccttctac 720  
 gccaccggcg acatcatcg cgacatccgc caggcccact gcaacatcg cggcgagaag 780  
 tgaacaaca ccctgaagca gatctgtacc aagctgcagg cccagttcgg caacaagacc 840  
 atcgtgttca agcagagcag cggcgccgac cccagatcg tgatgcacag cttcaactgc 900  
 ggcggcgagt tcttctactg caacagcacc cagctttca acagcacctg gaacaacacc 960  
 atcggccccca acaacacca cggcaccatc accctgcctt gccgcattca gcatatcatc 1020  
 aaccgctggg gcccgaaggc catgtacgccc ccccccattcc gcccgcagat cccgtcagc 1080  
 agcaacatca cccgcctgct gctgaccgc gacggcgca aggagatcg caacaccacc 1140  
 gagatcttcc gccccggcg cggcgacatg cggcacaact ggcgcagcga gctgtacaag 1200  
 tacaagggtgg tgaagatcg gccccctggc gtggccccc ccaaggccaa ggcgcgcgtg 1260  
 gtcagcgcg agaaggcgcc cgtgaccctg ggcgcattgt ttctgggctt cctggcgcc 1320  
 gccggcagca ccatggcgcc cccgcgttgc accctgaccg tgcaggcccg ccagctgtg 1380  
 agcggcattcg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcaagcacctg 1440  
 ctgcagctga cccgtgtggg catcaaggcag ctgcaggccc gcgtgtgc cgtgagcgc 1500  
 tacctgaagg accagcagct gctgggcattt tggggctgca gcccgaagct gatctgcacc 1560  
 accgcgtgc cctggaaacgc cagctggagc aacaagagcc tggaccagat ctggaaacaac 1620  
 atgaccttggg tggagtggg ggcgcagatc gacaactaca ccaacctgtt ctacaccctg 1680  
 atcgaggaga gcccgaacca gcaggagaag aacgagcagg agctgtgtt gctggacaag 1740  
 tggccagcc tggaaactg gttcgacatc agcaagtggc tggatcatc caagatcttc 1800  
 atcatgatcg tggccgcctt ggtggcctt cgcacatgtgt tcaccgtgt gaggatgtg 1860  
 aaccgcgtgc gccaggctt cagccccctt agcttccaga cccgcattcc cgcggccgc 1920  
 ggcggccgacc gccccgggg catcgaggag gagggcgccg agcgcgaccc cgaccgcagc 1980  
 agccccctgg tgcacggcctt gtcggccctt atctggacg acctgcgcag cctgtgcctg 2040  
 ttcagctacc accgcctgcg cgcacatgtt ctgatcgccg cccgcattgtt ggagctgtg 2100  
 ggcggccgcg gctggggccg cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160  
 gagctgaaga acagcgccgt gaggctgtt gacgcattatc ccatcgccgtt ggcggaggcc 2220  
 accgaccgcgca tcatcgaggat ggcggccgcg atcggccgcg ctttcctgca catccccccgc 2280  
 cgcatccgcgcc agggcttcga ggcgccttgc ctgtactcg ag 2322

&lt;210&gt; 21

&lt;211&gt; 2310

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Lys121-Val200;  
Asn425-Lys432

&lt;400&gt; 21

gaattcgccca ccatggatgc aatgaagaga gggctctgtt gttgtgtgtt gctgtgttgg 60  
 gcagtcttcg tttcgcccg cgcgcgtggag aagctgtggg tgaccgttta ctacggcggt 120  
 cccgtgtggg aggagccac caccaccctt ttctgcgcga ggcacgcctt ggccttacgc 180  
 accgaggtgc acaacgtgtt ggccacccac gcctgcgtgc ccaccgcacc caaccccccag 240  
 gagatgtgc tggagaacgtt gaccgagaac ttcaacatgtt ggaagaacaa catgtgtggag 300  
 cagatgcacg aggacatcat caccgtgtgg gaccgagacc tgaaggccctt cgtgaaggcc 360  
 cccgtgtatca cccaggcctt ccccaagggtt agcttcgagc ccatccccat ccactactgc 420  
 gccccccgcg gcttcgcctt cctgaagtgc aacgacaaga agttcaacgg cagggcccc 480  
 tgcaccaacg tgagcaccgtt gcaatgttgc accggcatcc gccccgtgtt gagcaccagg 540  
 ctgtgtgttgc acggcaccgtt ggccgaggag ggcgtgttgc tccgcagcga gaacttcacc 600  
 gacaacgcacca agaccatcat cgtgcacgtt aaggagagcg tggagatcaa ctgcacccgc 660  
 cccaacaaca acacccgcaaa gacatcacc atcggccccc gccgcgcctt ctacgcccacc 720  
 ggcgacatca tcggcgacat cccgcaggccc cactgcaaca tcaagcggcga gaagtggAAC 780  
 aacacccttgc agcagatgtt gaccaagctt caggcccattt tccgcacatc gaccatcg 840  
 ttcaaggcaga gcaaggccgg cgcaccccgatc atcgtgtatgc acagcttcaaa ctgcggccgc 900  
 gagttcttctt actgcacatc caccatcgatc ttcaacatcgat cctggaaacaa caccatcg 960  
 cccaacaaca ccaacgcac catcaccctt ccctgcgcga tcaagcagat catcaacgc 1020  
 cccaaggccca tgcacccccc cccatccgc ggcgcacatcc gctgcacgtt gacaccaccga 1080  
 ggcctgtgtgc tgaccgcga cggcgccgaag gagatcgatc acaccaccga gatcttccgc 1140  
 cccggcgccg ggcacatcgatc cgcacactgg cgcacgcac tgcataactcg caaggtgggt 1200

aagatcgagc ccctgggcgt ggcccccacc aaggccaagc gccgcgtggc 1260  
 aagcgcgcg tgaccctggg cgccatgttc ctgggcttcc tgggcgcgcg cggcagcacc 1320  
 atgggcggcc gcagcctgac cctgaccgtg caggccgcg agctgctgag cggcattgt 1380  
 cagcagcaga acaacctgtc gcgccatc gaggccgcg agcacctgtc gcagctgacc 1440  
 gtgtggggca tcaaggcgtc gcaggcccgcgtg cttggcttcc tggagcgtt 1500  
 cagcagctgc tggcattctg gggctgcgtc ggcaagctga tctgcaccac cggcgtgc 1560  
 tggAACGCCA gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620  
 gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgtat cgaggagagc 1680  
 cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtgc 1740  
 tggAACTGGT tcgacatcag caagtggctg tggatcatca agatcttcat catgatctgt 1800  
 ggcggcctgg tggcctgcgtc catcggttcc accgtgtga gcatcgtaa cggcgtgcgc 1860  
 cagggctaca gccccctgag ctccagacc cgcttccccg ccccccggg ccccgaccgc 1920  
 cccgagggca tcgaggagga gggcggcgag cgccggccgc accgcaggag cccctgggt 1980  
 cacggcctgc tggccctgtat ctgggacgac ctgcgcagcc tggcctgtt cagctaccac 2040  
 cgccctgcgcg acctgatcct gatcgccgcg cgcattgtgg agctgctggg cggccgcgc 2100  
 tggaggccc tgaagtactg gggcaacctg ctgcgtact ggttccaggaa gctgaagaac 2160  
 agcgcgtga gcctgttgcgtc cgccatcgcc atcggcgtgg ccgagggcac cgaccgcac 2220  
 atcgagggtgg cccagcgcat cgggccgcgc ttcctgcaca tccccggccg catccgcag 2280  
 ggcttcgagc ggcgcctgtc gtaactcgag 2310

&lt;210&gt; 22

&lt;211&gt; 2298

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Val120-Ile201;  
Ile424-Ala433

&lt;400&gt; 22

gaattcggcca ccatggatgc aatgaagaga gggctctgtgtc gtgtgtgg 60  
 gcagtcttcg ttccggccag cgccgtggag aagctgtggg tgaccgtgtt ctacggcgtg 120  
 cccgtgttgg aaggaggccac caccaccctg ttctgcgcga ggcacgcac 180  
 accgagggtgc acaacgtgtg gggcaccac gcctgcgtgc ccaccgaccc 240  
 gagatcgatgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctgt cgtggcggc 360  
 atcaccctagg cctggcccaa ggtgagcttcc gggccatcc ccatccacta ctgcgc 420  
 gcccgttccg ccattctgaa gtgcacgcac aagaaggta acggcagcg 480  
 aacgtgagca cctgtcgttgc caccacggc atccgc 540  
 ctgaacggca ggcctggccgaa ggaggccgtg gtatccgc gcgagaactt caccgacaac 600  
 gccaagacca tcatcgatgc gctgaaggag agcgtggaga tcaactgcac cggcccaac 660  
 aacaacaccc gcaagagcat caccatggc cccggccgcg ctttctacgc caccggc 720  
 atcatcgccg acatccgcca gggccacttgc aacatcagcg gcgagaactt gaaacaacacc 780  
 ctgaagcaga tgcgtacccaa gctgcaggcc cagttccgc acaagaccat cgtgttcaag 840  
 cagagcagcg gcccgcgcaccc cgagatcgatgc atgcacacat tcaactgcgcg cggcgcgtt 900  
 ttctactgtca acagcacccaa gctgttcaac agcacctggaa acaacacccat cggcccaac 960  
 aacaccaacg gcaccatcac cctggccctgc cgcatcaagc agatcatcg 1020  
 tacggccccc ccattccgcg ccagatccgc tgcagcagca acatcaccgg cctgtgtctg 1080  
 accccgcacg gcccgcacca gatcagcaac accaccggaa tcttccgc 1140  
 gacatcgccg acaactggcg cagcgagctg tacaaggata agtgggtgaa gatcgagccc 1200  
 ctgggcgtgg ccccccacca ggcacgcgc cgcgtggcgtc agcgcgcgaa ggcgcgcgtg 1260  
 accctggccg ccatgttccct gggcttccgt ggcgcgcgcg gcagcaccat gggcccccgc 1320  
 agcctgaccc tgaccgtgc gggccgcgcg ctgctgagcg gcatcgatgc gcagcagaac 1380  
 aacctgctgc ggcacgcgc gggccgcgcg cacctgtgc agtgcaccgt gtgggcattc 1440  
 aagcagctgc aggcccgcgt gctggccgtg gagcgttacc tgaaggacca gcagctgt 1500  
 ggcacatctggg gctgcagcgaa caagctgtatc tgcaccaccgc cctgtgc 1560  
 tggagcaaca agagcctggaa ccagatctgg aacaacatgtc cttggatggaa gtggagcgc 1620  
 gagatcgaca actacacccaa cctgtatctac accctgatgc aggagagccaa gaaccagcag 1680  
 gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggatc 1740

gacatcagca agtggctgtg gtacatcaag atttcatca tgatcgtggg cggcctggg 1800  
ggcctgcgca tcgtgttac cgtgctgagc atcgtgaacc gcgtgcgcac gggctacagc 1860  
cccctgagct tccagacccg ctccccgc ccccgccgccc cccgaccgc cgggggcattc 1920  
gaggaggagg gggcgagcg cgaccgcac cggcagcagcc ccctggtgc a cggcctgctg 1980  
gccctgatct gggacgaccc ggcgcagcctg tgccctgttca gctaccacccg cctgcgcgac 2040  
ctgatcctga tcgcccggcc catcgtggag ctgctgggccc gccgcggctg ggaggccctg 2100  
aagtactggg gcaacctgct gcagttactgg atccaggagc tgaagaacag cgccgtgagc 2160  
ctgttcgacg ccatcgccat cggcgtggcc gagggcaccg accgcacatcat cgaggtggcc 2220  
cagcgcacatcg gcccgcctt cctgcacatc ccccgccgca tccgcacaggg cttcgagcgc 2280  
qccctgctgt aactcgag 2298

<210> 23

<211> 2298

<212> DNA

<213> Artificial Sequence

<220>

**<223> Description of Artificial Sequence:**

Val120-Ile201B: Ile424-Ala433

<400> 23

gccctgctgt aactcgag

2298

&lt;210&gt; 24

&lt;211&gt; 2298

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Val120-Thr202;  
Ile424-Ala433

&lt;400&gt; 24

gaattcggcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtcttcg ttgcggccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180  
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccacccgaccc caaccccccag 240  
 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtgag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggccggc 360  
 gcccacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcggcc 420  
 gccggcttcg ccatacctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480  
 aacgtgagca ccgtgcagtg caccacggc atccggcccg tggtgagcac ccagctgctg 540  
 ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600  
 gccaagacca tcatacgtgca gctgaaggag agcgtggaga tcaactgcac cggccccaac 660  
 aacaacaccc gcaagagcat caccatcgcc cccggcccg ccttctacgc caccggcgac 720  
 atcatcgccg acatccggca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780  
 ctgaagcaga tcgtgaccaa gctgcaggcc cagttcgca acaagaccat cgtgttcaag 840  
 cagagcagcg gcgccgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagtgc 900  
 ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960  
 aacaccaacg gcaccatcac cctgccttcg cgcatacgc agatcatcg cggcgccatg 1020  
 tacggccccc ccataccggc ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080  
 acccgcacg gcgccaaagga gatcagcaac accacccgaga tcttccggcc cggcgccggc 1140  
 gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200  
 ctgggcgtgg ccccccacca ggccaagcgc cgcgtggtgc agcgcgagaa gcgccgcgtg 1260  
 accctggcgcc ccatgttccctt gggcttcctg ggccggcccg gcagcaccat gggcccccgc 1320  
 agcctgaccc tgaccgtgca ggcccggccag ctgctgagcg gcatcgtgca gcagcagaac 1380  
 aacctgtgc gcgcacatcgaa ggcccagcag cacctgtgc agctgaccgt gtggggcatc 1440  
 aagcagctgc agggcccggt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500  
 ggcatactggg gtcgcagcgg caagctgatc tgcaccaccg cgcgtccctg gaacgcacgc 1560  
 tggagcaaca agagccttgg ccagatctgg aacaacatga cctggatggc gtggagcgc 1620  
 gagatcgaca actacaccaa cctgatctac accctgatcg aggagagccca gaaccagcag 1680  
 gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740  
 gacatcagca agtggctgtg gtacatcaag atcttcatca tgcgtgggg cggcctggtg 1800  
 ggcctgcgca tcgtgttccac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacage 1860  
 cccctgagct tccagaccccg cttcccccggc ccccgccggcc cgcaccggcc cgagggcatc 1920  
 gagggaggagg gcgccgagcgc cgaccgcgc cgcagcagcc cctgtgtca cggcctgtg 1980  
 gcccctgatct gggacgaccc ggcgcggcc cgcgcggcc cgcgtggccgc gtcaccaccg cctgcgcgc 2040  
 ctgatcctga tcgcgcggcc catcggtggag ctgctggcc gcccggcgtg ggaggccctg 2100  
 aagtactggg gcaacctgtc gcagtaactgg atccaggagc tgaagaacag cgcgcgtg 2160  
 ctgttcgcacg ccatacgcacat cgcgcgtggcc gagggcaccgc accgcacatcat cgcggccggc 2220  
 cagcgcacatcg gccgcgcctt cctgcacatc ccccgccgc tccgcacaggc cttcgagcgc 2280  
 gcccctgctgt aactcgag 2298

&lt;210&gt; 25

&lt;211&gt; 2358

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val127-Asn195

&lt;400&gt; 25

gaattcggcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtcttcg tttcgccca ggcgcgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccacccctg ttctgcggca ggcgcgtggac 180  
 accgaggtgc acaacgtgtg ggcacccac gcctgcgtgc ccacccgaccc caaccccccag 240  
 gagatgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 acccccccgt gcgtggggc aggaaactgc aacaccagcg tgatcaccacca ggcctgcccc 420  
 aaggtagct tcgagccat ccccatccac tactgcgccc cgcgcgtt cgcgcgtt 480  
 aagtgcacg acaagaagtt caacggcgc ggcgcgtca ccaacgtgag caccgtgcag 540  
 tgcacccacg gcatccggcc cgtgggtgagc acccagctgc tgctgaacgg cagcctggcc 600  
 gaggaggggcg tggtgatccg cagcgagaac ttcaaccgaca acgccaagac catcatcg 660  
 cagctgaagg agagcgtgg aatcaactgc accccgcacca acaacaacac cgcgaagagc 720  
 atcaccatcg gccccggccg cgccctctac gccaccggcg acatcatcg cgcacatccgc 780  
 caggcccact gcaacatcatc cggcgagaag tggaaacaaca ccctgaagca gatcggtgacc 840  
 aagctgcagg cccagttcgg caacaagacc atcggttca agcagagcag cggcggcgac 900  
 cccgagatcg tgatgcacag cttcaactgc ggcggcgagt tcttctactg caacagcacc 960  
 cagctgttca acagcacctg gaacaacacc atcgcccca acaacaccaa cggcaccatc 1020  
 accctgcctt gccgcatcaa gcagatcatc aaccgctggc aggaggtggg caaggccatg 1080  
 tacgcccccc ccatccgcgg ccaagatccgc tgcaagcagca acatcaccgg cctgctgtg 1140  
 accccgcacg gcgcaagga gatcagcaac accaccgaga tcttccgccc cggcggcgcc 1200  
 gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtaa gatcgagccc 1260  
 ctggcggtgg ccccccaccaa ggcacagcgc cgcgtggc agcgcgagaa ggcgcgcgtg 1320  
 accctggcg ccatgttctt gggcttctg ggcgcgcgg gcagcaccat gggcgcgcgc 1380  
 agcctgaccc tgaccgtgca ggcgcgcag ctgctgagcg gcatcggtca gcaagcagaac 1440  
 aacctgctgc ggcacatcga ggcgcagcag cacctgctgc agctgaccgt gtggggcatc 1500  
 aagcagctgc aggccccgt gctggccgtg gagcgctacc tgaaggacca gcaagctgtg 1560  
 ggcacatctgg gctgcagcgg caagctgatc tgcaaccaccg ccgtgcgcgtg gaacgcgc 1620  
 tggagcaaca agagccttgg ccaagatctgg aacaacatga cctggatggg gtgggagcgc 1680  
 gagatcgaca actacaccaa cctgatctac accctgatcg aggagagcca gaaccagcag 1740  
 gagaagaacg agcaggagct gctggagctg gacaagtggg ccagecgttg gaaactgtt 1800  
 gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcggtgg cggcctggc 1860  
 ggcctgcgca tcgtgttccac cgtgtgtgac atcggttca ggcgcgcac gggctacagc 1920  
 cccctgagct tccagacccg cttcccccgc cccgcggcc ccgaccgcgc cggggcatac 1980  
 gaggaggagg gcgccgagcg cgcgcgcac cgcagcagcc ccctggtgca cggcctgtg 2040  
 gcccgtatct gggacgaccc ggcgcgcgtg tgccgttca gctaccaccg cctgcgcgc 2100  
 ctgatcctga tcgcgcgcgc catcggtggag ctgctggcc ggcgcggctg ggaggccctg 2160  
 aagtactggg gcaacctgtc gcagtaactgg atccaggagc tgaagaacag cgcgcgtgac 2220  
 ctgttcgacg ccatcgccat cgcgcgtggc gagggcaccg accgcacatc cggggcggcc 2280  
 cagcgcacatc ggcgcgcctt cctgcacatc cccgcgcac tccgcacagg cttcgacgc 2340  
 gcccgtgt aactcgag 2358

&lt;210&gt; 26

&lt;211&gt; 2352

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Val127-Asn195;  
Arg426-Gly431

&lt;400&gt; 26

gaattcggcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60  
 gcagtcttcg tttcgccca ggcgcgtggag aagctgtggg tgaccgtgta ctacggcgtg 120  
 cccgtgtgga aggaggccac caccacccctg ttctgcggca ggcgcgtggac 180  
 accgaggtgc acaacgtgtg ggcacccac gcctgcgtgc ccacccgaccc caaccccccag 240  
 gagatgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360  
 acccccccgt gcgtggggc aggaaactgc aacaccagcg tgatcaccacca ggcctgcccc 420

aaggtagct tcgagccat ccccatccac tactgcgcc cggccggctt cgccatcctg 480  
aagtgcacg acaagaagtt caacggcagc ggccctgcac ccaacgtgag caccgtgcag 540  
tgcacccacg gcatccccc cgtggtgagc acccagctgc tgctgaacgg cagcctggcc 600  
gaggaggcg tggtgatccg cagcgagaac ttcaccgaca acgccaagac catcatcg 660  
cagctgaagg agagcgtgga gatcaactgc acccgccccca acaacaacac ccgcaagagc 720  
atcaccatcg gccccggccg cgccctctac gccaccggcg acatcatcg ccacatccgc 780  
caggcccact gcaacatcg cggcgagaag tggacaaca cctgaaagca gatcgtgacc 840  
aagctgcagg cccagttcg caacaagacc atcgtgttca agcagagcag cggcggcgc 900  
cccgagatcg ttagtgcacag cttaactgc ggcggcgagt tcttctactg caacagcacc 960  
cagctgttca acagcacctg gaacaacacc atcgccccca acaacacccaa cggcaccatc 1020  
accctgcctt gccgcataa gcagatcatc aaccgcggcg gcccgaaggc catgtacgcc 1080  
ccccccatcc gcccggcagat ccgctgcage agcaacatca cccgcctgtc gctgaccgc 1140  
gacggcggca aggagatcg caacaccacc gagatcttc gccccggggg cggcgacatg 1200  
cgcgacaact ggcgcagcga gctgtacaag tacaagggtgg tgaagatcga gcccctggc 1260  
gtggccccc ccaaggccaa gcccgcgtg gtgcagcgcg agaagcgcgc cgtgaccctg 1320  
ggcgccatgt tcctgggtt cctggcgcc gccggcagca ccatgggcgc ccgcagcctg 1380  
accctgaccg tgcaggcccc ctagctgctg agcggcatcg tgcagcagca gaacaacctg 1440  
ctgcgcgcca tgcaggcccc gcagcacctg ctgcagctga ccgtgtgggg catcaagcag 1500  
ctgcaggccc gcgtgtcgcc cgtggagcgc tacctgaagg accagcagct gctggcattc 1560  
tggggctgca gcccggcagat gatctgcacc accggcgtgc cctggAACgc cagctggagc 1620  
aacaagagcc tggaccagat ctggaacaac atgaccttgc tggagtggga gcgcgagatc 1680  
gacaactaca ccaacctgat ctacaccctg atcgaggaga gccagaacca gcaggagaag 1740  
aacgagcagg agctgttgc gctggacaag tggggcagcc tggactgc gttcgacatc 1800  
agcaagtggc tgggtacat caagatctt atcatgatcg tggggcgctt ggtggccctg 1860  
cgcatcggt tcaccgtgt gacatcgat aaccgcgtgc gcccggctt cagccccctg 1920  
agcttccaga cccgccttccc cgcggggccg gggccggacc gcccggaggg catcgaggag 1980  
gagggcggcg agcgcgaccg cgaccgcagc agccccctgg tgcacggctt gctggccctg 2040  
atctggacg acctgcgcag cctgtgcctg ttcaactacc accgcctgcg cgacctgtatc 2100  
ctgatcgccg cccgcatacg tggactgtgc ggcgcggcg gctggggaggc cctgaagtac 2160  
tggggcaacc tgctgcagta ctggatccag gagctgaaga acagcgcgtt gacccctgtt 2220  
gacccatcg ccatacgccgt ggcgcggggc accgaccgc tcatcgaggt ggcggccgtt 2280  
atcgccgcgc cttccctgca catccccccgc cgcatccgc accggcttcga ggcgcgcctg 2340  
ctgtactcg ag 2352